

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:09 ; Search time 76 Seconds

(Without alignments)
1002.886 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061

Sequence: 1 RESLCGVVFLLAAGLPIQ.....PFSSGDKREKPLQKPKWML 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing First 45 summaries

Database :

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3040	99.3	572	19	AAK38335
2	2700	88.2	574	21	AA15119
3	2163	70.7	560	18	AAK35382
4	2163	70.7	560	23	AA15119
5	2163	70.7	560	23	AA15119
6	2163	70.7	560	23	AA15119
7	644.5	21.1	376	18	AAK35386
8	644.5	21.1	376	18	AAK35387
9	642	21.0	625	20	AA131978
10	627.5	20.5	661	16	AAK78646

11	627.5	20.5	661	20	AA131977	Human melanoma ant
12	627.5	20.5	661	22	AA131978	Human melanoma ant
13	627.5	20.5	661	22	AA131979	Melanoma antigen C
14	627.5	20.5	661	23	AAU29003	Human gp100 consen
15	627.5	20.5	661	23	AAU84803	Human melanoma ant
16	627.5	20.5	661	23	AAU09695	Human melanoma ant
17	626.5	20.5	661	16	AA131978	MA17-1 melanoma an
18	626.5	20.5	661	16	AA131979	Human melanoma ant
19	624.5	20.4	661	22	AA131978	Melanoma antigen C
20	624.5	20.4	661	22	AA131979	Human melanoma ant
21	622.5	20.3	661	16	AA131978	MA17-1 melanoma an
22	622.5	20.3	661	16	AA131979	Modified tumour-as
23	622.5	20.3	661	22	AA131978	Human gp100M prote
24	622.5	20.3	661	22	AA131979	Modified gp100M pr
25	619.5	20.2	661	22	AA131978	Human melanoma ant
26	619.5	20.2	661	22	AA131979	Human melanoma ant
27	616.5	20.1	661	22	AA131978	Human melanoma ant
28	615.5	20.1	661	22	AA131979	Human melanoma ant
29	614.5	20.1	661	22	AA131978	Human melanoma ant
30	614.5	20.1	661	22	AA131979	Human melanoma ant
31	613.5	20.1	661	22	AA131978	Human melanoma ant
32	612.5	15.1	105	22	AA131979	Peptide #1551 enco
33	612.5	15.1	105	22	AA131978	Peptide #1551 enco
34	612.5	15.1	105	22	AA131979	Peptide #1551 enco
35	612.5	15.1	105	22	AA131978	Human brain expres
36	612.5	15.1	105	22	AA131979	Human bone marrow
37	612.5	15.1	105	22	AA131978	Peptide #1551 enco
38	612.5	15.1	105	22	AA131979	Peptide #1551 enco
39	612.5	15.1	105	22	AA131978	Peptide #1551 enco
40	612.5	15.1	105	22	AA131979	Human peptide enco
41	295	9.6	69	22	AA131978	Peptide #11493 enco
42	295	9.6	69	22	AA131979	Protein #8895 enco
43	295	9.6	69	22	AA131978	Protein #9161 enco
44	295	9.6	69	22	AA131979	Human melanoma ant
45	295	9.6	69	22	AA131978	Human bone marrow

ALIGNMENTS

RESULT 1
ID AAK38335 standard; Protein: 572 AA.

AAK38335
AC AAK38335;
AD 21-MAY-1998 (first entry)
DE Rat kidney injury related molecule (KIM).

KW Kidney injury related molecule; KIM; rat; renal disease; injury;
KW nephritic; tissue regeneration; therapy; monoclonal antibody.

OS Rattus sp.

PN WO9744460-A1.

PD 27-NOV-1997.

XX 23-MAY-1997; 97WO-US09303.

XX 23-AUG-1996; 96US-0023442.

XX 24-MAY-1996; 96US-0018228.

XX (BIO) EIOGEN INC.

XX Bonventre JV, Cate RL, Hession CA, Ichimura T, Sanicola-Nadel M;

XX Wei H;

XX WPI; 1998-018514/02.

XX DR N-PSDB: AAT96034.

XX DNA encoding kidney injury related molecule - which is upregulated

PT in injured or regenerating tissue, useful to promote growth of new
 XX tissue and survival of damaged tissue
 PS Claim 9; Page 41-43; 68pp; English.

XX This protein, designated kidney injury related molecule (KIM), is
 CC up-regulated in injured or regenerating tissue, is an amino acid
 CC chain, and is secreted from cells, including cells from
 CC ischemic adult rat kidneys. A 307-amino acid rat KIM (see
 CC AAM38334) and a human KIM (see AAM36336) are also claimed. Recombinant
 CC KIM polypeptides can be expressed in prokaryotic and eukaryotic
 CC host cells using a claimed process. Soluble variants fused to
 CC a toxin, aneugenic compound or radionuclide, and IgG fusion
 CC proteins are also claimed. The KIM polypeptides can be used to treat
 CC renal disease and to promote the growth of new tissue or the
 CC survival of damaged tissue, generally in conditions where the
 CC binding of specific ligand to KIM stimulates cell growth, maintains
 CC cellular differentiation or reduces apoptosis, e.g. in cases of
 CC renal failure, nephritis, kidney transplants, toxic or hypoxic
 CC injury. A monoclonal antibody specific for KIM can be used to
 CC detect and measure KIM in biological fluids. KIM is involved in
 CC neoplasia, loss of cellular function, susceptibility to apoptosis
 CC or promotion of inflammation, deliver imaging agents to KIM
 CC expressing cells in vivo or in vitro and measure KIM concentration
 CC by immunoassay. Damage/regeneration of renal cells can be determined
 CC by measuring KIM, particularly to diagnose or monitor the progress
 CC of disease or therapy. KIM-expressing tumour cells can be inhibited
 CC by KIM, and KIM can be used to deliver cytotoxic agents, such as
 CC a toxin or radionuclide, and tumour cells that express KIM ligand
 CC can be inhibited with similarly tagged KIM or anti-KIM ligand
 CC antibody.

XX Sequence 572 AA:

Query Match 99.38; Score 3040; DB 19; Length 572;
 Best Local Similarity 99.38; Pred. No. 7.1e-272;
 Matches 568; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESICGVYVLLAAGLPQAARFPRDVLGHEQYPDHMRNNQLKNGSSDENEDQLYP 60
 DB 1 MESICGVYVLLAAGLPQAARFPRDVLGHEQYPDHMRNNQLKNGSSDENEDQLYP 60
 QY 61 VRRKGGKRWKDSNKGGRVQAAITSDSPALVGSNITVYVNLVPRCKDANGNHYVERNC 120
 DB 61 VRRKGGKRWKDSNKGGRVQAAITSDSPALVGSNITVYVNLVPRCKDANGNHYVERNC 120
 QY 121 RSDELASDPYVYNTGADDEWNTSQGHLRFQDQPPRPHGRKKNVYVPHTL 180
 DB 121 RSDELASDPYVYNTGADDEWNTSQGHLRFQDQPPRPHGRKKNVYVPHTL 180
 QY 181 GYVFKRLQGCARSYNTVNLVYGVQVMEVTVFRHGRAYIPISKVDVYVITDQIPFV 240
 DB 181 GYVFKRLQGCARSYNTVNLVYGVQVMEVTVFRHGRAYIPISKVDVYVITDQIPFV 240
 QY 241 TMVKQKRNDSNENRFGRLDPIFPDVLHDSHLYANSAIYKWNFGDNTGLVYSNHTIL 300
 DB 241 TMVKQKRNDSNENRFGRLDPIFPDVLHDSHLYANSAIYKWNFGDNTGLVYSNHTIL 300
 QY 301 HTVYLVNGTFNPLVTQVATVAGCPSPFSPSSSTSPSPASSSTPTLSPSPSLMPTGYKS 360
 DB 301 HTVYLVNGTFNPLVTQVATVAGCPSPFSPSSSTSPSPASSSTPTLSPSPSLMPTGYKS 360
 QY 361 MELSDISNENCRINRGYFRATITVDGLVLEVINIQTADVPIPTLOPNSLMDQFVTCGK 420
 DB 361 MELSDISNENCRINRGYFRATITVDGLVLEVINIQTADVPIPTLOPNSLMDQFVTCGK 420
 QY 421 ATPFACATITSDPFOQIAQRNQSVPANDEICLLSVRRANNGSGTCYCNVFTLGDASIAL 480
 DB 421 ATPFACATITSDPFOQIAQRNQSVPANDEICLLSVRRANNGSGTCYCNVFTLGDASIAL 480
 QY 481 TSALISIPGKDGSPRLTVNGVLISIGCLAMFVMTVITLLYKKHYTKPGNCTRNVRGK 540
 DB 481 TSALISIPGKDGSPRLTVNGVLISIGCLAMFVMTVITLLYKKHYTKPGNCTRNVRGK 540

OY 541 KGLSVFLSHAKAFPSRGDRKDEKDPILLQDKPWML 572
 DB 541 KGLSVFLSHAKAFPSRGDRKDEKDPILLQDKPWML 572

RESULT 2

AA19119
 ID AAB19119 standard; protein: 574 AA.

XX AAB19119;
 AC AAB19119;
 XX 19-FEB-2001 (first entry)
 DT 19-FEB-2001 (first entry)
 DE Polypeptide isolated from lymph node stromal cells of fsn +/- mice.
 XX Lymph node stromal cell; fsn +/- mice: inflammatory disorder;
 KW immune system growth; cancer; viral disorder; HIV infection;
 KW blood vessel growth; tumor necrosis factor disorder; arthritis;
 KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;
 XX Cardiac failure.

OS Mus sp.

XX W0200058463-A1.

XX 05-OCT-2000.

XX 18-FEB-2000; 2000WO-NZ00015.

XX 25-MAR-1999; 99US-0276268.

XX 26-AUG-1999; 99US-0383586.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;

XX Murison JG;

XX WPI: 2000-564924/64.

XX N-PSDB; AAM56729.

XX polypeptide expressed in mammalian fsn +/- lymph node stromal cells,
 XX useful for modulating growth of blood cells, for treating inflammatory
 XX and tumour necrosis factor-mediated disorders, cancer and viral
 XX disorders

XX Claim 1; Page 54-55; 75pp; English.

XX The present sequence represents a polypeptide sequence which is
 CC isolated from lymph node stromal cells of fsn +/- mice. The
 CC polynucleotides and their polypeptides are useful for treating an
 CC inflammatory disorder, disorder of immune system and cancer selected
 CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a
 CC viral disorder, inflammatory bowel disease, fibroblast growth factor-mediated
 CC disorder, immune system growth, cancer, viral infection, HIV infection,
 CC tumour necrosis factor (TNF) mediated disorder, such as those selected
 CC from arthritis, inflammatory bowel disease and cardiac failure and a
 CC fibroblast growth factor-mediated disorder. It is also useful in assays
 CC to determine biological activity, to raise antibodies, to isolate
 CC corresponding ligands or receptors, to quantify levels of protein or
 CC nucleic acid, to identify and clone genes and to identify amino acids
 CC and in compositions for the treatment of skin, connective tissue and
 CC immune system diseases. The polynucleotide is useful as marker for
 CC tissue, as a chromosome marker or tags in the identification of a
 CC genetic disorder.

XX Sequence 574 AA:

Query Match 88.24; Score 2700; DB 21; Length 574;
 Best Local Similarity 88.84; Pred. No. 1.9e-240;
 Matches 507; Conservative 20; Mismatches 42; Indels 2; Gaps 1;

OY 1 MESICGVYVLLAAGLPQAARFPRDVLGHEQYPDHMRNNQLKNGSSDENEDQLYP 60

[illegible]

[illegible]

Db 241 TWFOKDRNSDPTLKDLPIMFDVLIHDSHFLLNSTYNTKMSFGDNTGLFVSTNHTVN 300
 Qy 301 HTYVLNGTFNNLTWOTAVGPGDPSPTSSSTSPSPASSPTLST-----PSLSLMP 355
 Db 301 HTYVLNGTFNNLTWOTAAAGPGCPDPPPPR-----PSKPTPSLATILKSYDSNTGCP 353
 Qy 356 TGYSKMSLSDISENCRNKGTYGFRATITVVGILEVNTIOVADPIPTLOPNSLMDFI 415
 Db 356 TGYSKMSLSDISENCRNKGTYGFRATITVVGILEVNTIOVADPIPTLOPNSLMDFI 415
 Qy 354 TDNLESLRIPDPCNQIRHQIQAATITVVGILEVNTIOVADPIPTLOPNSLMDFI 413
 Db 354 TDNLESLRIPDPCNQIRHQIQAATITVVGILEVNTIOVADPIPTLOPNSLMDFI 413
 Qy 416 VTCGATPACTACTCTCOLAQNRCVSPVAVDELCLISVBRFNGSGTYCVNFTLGGD 475
 Db 416 VTCGATPACTACTCTCOLAQNRCVSPVAVDELCLISVBRFNGSGTYCVNFTLGGD 475
 Qy 476 ASIALTSALISIPGRDGLSGPRLTYVNGVLSIGCLAMFYVWITILLYKKHRTYKPGINCTR 535
 Db 476 ASIALTSALISIPGRDGLSGPRLTYVNGVLSIGCLAMFYVWITILLYKKHRTYKPGINCTR 535
 Qy 536 NVYKGLSVFLNARAKAVFFPGNOEKDPLKQ 568
 Db 536 NVYKGLSVFLNARAKAVFFPGNOEKDPLKQ 568
 Qy 534 NVYKGLSVFLNARAKAVFFPGNOEKDPLKQ 566
 Db 534 NVYKGLSVFLNARAKAVFFPGNOEKDPLKQ 566

RESULT 8
 AAW35386
 ID AAW35386 standard; Protein; 376 AA.
 XX
 AC AAW35386;
 XX
 DT 26-FEB-1998 (first entry)
 DE Murine metastatic nucleic acid sequence product.
 KW Mouse; murine; tumour; cancer; metastatic sequence; detection;
 KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;
 KW hypertrophy; screening.
 XX
 XX MMS musculus.
 XX
 XX W09718454-A2.
 XX
 XX 22-MAY-1997.
 XX
 XX 15-NOV-1996; 96WO-US18567.
 XX
 XX 30-JAN-1996; 96US-0594031.
 XX
 XX 16-NOV-1995; 95US-0006838.
 XX
 XX (THOM/) THOMPSON T.
 XX
 XX Thompson T;
 XX
 XX WPI: 1997-289397/26. .
 XX
 XX Identifying tumour metastatic sequences - by introducing transfected
 XX cells into host mammal and analysing primary and metastatic
 XX sequences by differential display PCR
 XX
 XX Disclosure; Fig 12CR: 102pp; English.
 XX
 XX Mouse urogenital sinus (US) tissue was isolated from 17 day old
 XX mouse embryos. The US cells were isolated with collagenase,
 XX cultured and implanted under the renal capsule of mice.
 XX Reconstitutions were harvested 5 weeks later, when they showed
 XX signs of distress from the tumour burden. Metastasised tumours were
 XX isolated from a site outside the renal capsule. RNA was isolated
 XX from primary tumours and metastases, reverse transcribed and
 XX subjected to differential display PCR. The sequences were analysed
 XX and compared with the sequences of the endogenous genes.
 XX present sequence. The method can be used to detect diagnosis and
 XX treat disorders related to metastasis, or treat malignant or
 XX non-malignant disorders, e.g. hyperplasia, dysplasia and
 XX hypertrophy. The metastatic sequence can be used to screen a

CC biological sample for metastasis, and it or its expression product
 CC may also be used to treat a metastatic disorder.
 XX
 XX Sequence 376 AA;
 XX
 XX Query Match 21.1%; Score 644.5; DB 18; Length 376;
 XX Best Local Similarity 38.9%; Pred. No. 1.1e-50;
 XX Matches 213; Conservative 54; Mismatches 101; Indels 179; Gaps 47;
 QY 16 GLPLQAKRFRVYLGHEQYFDHNRNQLAGKNSDNRNKGDLFPVRRGSGNSWKG 75
 Db 16 GLPLQAKRFRVYLGHEQYFDHNRNQLAGKNSDNRNKGDLFPVRRGSGNSWKG 75
 QY 5 GAARDAAR-RHDVGNRSAY---MRHNN---GWSSD-NOMNRY---VVKRGDMRWKNSKG 53
 Db 5 GAARDAAR-RHDVGNRSAY---MRHNN---GWSSD-NOMNRY---VVKRGDMRWKNSKG 53
 QY 76 GRVQAALTSIPALVNSITFVYVNLVFPCKEDANGNIVYERNCRSDIELASDPYVNN 135
 Db 76 GRVQAALTSIPALVNSITFVYVNLVFPCKEDANGNIVYERNCRSDIELASDPYVNN 135
 QY 136 TRG-ADDESMDNMTSOQHLPFRFPFRGRKRYKNEVYFHTLGOYFQKLGCSARNS 195
 Db 136 TRG-ADDESMDNMTSOQHLPFRFPFRGRKRYKNEVYFHTLGOYFQKLGCSARNS 195
 QY 97 TANSND---SDNGTSGSHNY-DGR---HGWRRNNTYHTGYK-----GRCSVRNS 139
 Db 97 TANSND---SDNGTSGSHNY-DGR---HGWRRNNTYHTGYK-----GRCSVRNS 139
 QY 196 INTVNLTVGPMVEIVFRRHGRAYIPISKVDYVITDQIPIPTVMTOKDRNSDPT 255
 Db 196 INTVNLTVGPMVEIVFRRHGRAYIPISKVDYVITDQIPIPTVMTOKDRNSDPT 255
 QY 140 NTANNT---GHVTVYRIRGRAY---AYRDYVYVD---VYTH-KADRNSQTK- 184
 Db 140 NTANNT---GHVTVYRIRGRAY---AYRDYVYVD---VYTH-KADRNSQTK- 184
 QY 256 LRDLPIPPDVLHDPSPUNYSAISYKWNFGDNTGLFVSNHNTLHNTYVNGTTFNLT 315
 Db 256 LRDLPIPPDVLHDPSPUNYSAISYKWNFGDNTGLFVSNHNTLHNTYVNGTTFNLT 315
 QY 185 -----DNDVHD-SH-NYST-NYKWS-GDNTG--VSTNHTVHTYV-NCTSN--TV 225
 Db 185 -----DNDVHD-SH-NYST-NYKWS-GDNTG--VSTNHTVHTYV-NCTSN--TV 225
 QY 316 QTAVPGPSPPTSPSSSTSPSPASSPTLSTPSPSPSPSPSPSPSPSPSPSPSPSP 375
 Db 316 QTAVPGPSPPTSPSSSTSPSPASSPTLSTPSPSPSPSPSPSPSPSPSPSPSPSP 375
 QY 376 YGYFRATITVVGILEVNTIOVADPIPTLOPNSLMDFIPTCKGATPTEACTISDPTC 435
 Db 376 YGYFRATITVVGILEVNTIOVADPIPTLOPNSLMDFIPTCKGATPTEACTISDPTC 435
 QY 249 TGH-----ATTVG-----VNMTDVMWSSD-----VYTC-----GSTVCTSDTC 282
 Db 249 TGH-----ATTVG-----VNMTDVMWSSD-----VYTC-----GSTVCTSDTC 282
 QY 436 QIACNRVCSFVAVDELCLISVBRFNGSGTYCVNFTLGDASIALTSALISIPGRDGLSP 495
 Db 436 QIACNRVCSFVAVDELCLISVBRFNGSGTYCVNFTLGDASIALTSALISIPGRDGLSP 495
 QY 283 ---TNTVCS-VYVD---NCTVRRP-NQSTTCVYV---TDDTTSATSTV---DRDAS-- 325
 Db 283 ---TNTVCS-VYVD---NCTVRRP-NQSTTCVYV---TDDTTSATSTV---DRDAS-- 325
 QY 496 LRTVNCVLISIGCLAMFYVWITILLYKKHRTYKPGINCTRNVYKGLSVFLSHAKAPFS 555
 Db 496 LRTVNCVLISIGCLAMFYVWITILLYKKHRTYKPGINCTRNVYKGLSVFLSHAKAPFS 555
 QY 326 ---RWANSA---SVGC-----AVTVSVYKHKY-----NNSGNVYRSGK---SVNRKAY-- 366
 Db 326 ---RWANSA---SVGC-----AVTVSVYKHKY-----NNSGNVYRSGK---SVNRKAY-- 366
 QY 556 RDRKSD 562
 Db 556 RDRKSD 562
 QY 367 -GNKDKN 372
 Db 367 -GNKDKN 372

RESULT 8
 AAW35387
 ID AAW35387 standard; Protein; 376 AA.
 XX
 AC AAW35387;
 XX
 DT 26-FEB-1998 (first entry)
 DE Murine metastatic nucleic acid sequence product.
 KW Mouse; murine; tumour; cancer; metastatic sequence; detection;
 KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;
 KW hypertrophy; screening.
 XX
 XX MMS musculus.
 XX
 XX W09718454-A2.
 XX
 XX 22-MAY-1997.
 XX
 XX 15-NOV-1996; 96WO-US18567.
 XX
 XX 30-JAN-1996; 96US-0594031.
 XX
 XX 16-NOV-1995; 95US-0006838.
 XX
 XX (THOM/) THOMPSON T.
 XX
 XX Thompson T;
 XX
 XX WPI: 1997-289397/26. .
 XX
 XX Identifying tumour metastatic sequences - by introducing transfected
 XX cells into host mammal and analysing primary and metastatic
 XX sequences by differential display PCR
 XX
 XX Disclosure; Fig 12CR: 102pp; English.
 XX
 XX Mouse urogenital sinus (US) tissue was isolated from 17 day old
 XX mouse embryos. The US cells were isolated with collagenase,
 XX cultured and implanted under the renal capsule of mice.
 XX Reconstitutions were harvested 5 weeks later, when they showed
 XX signs of distress from the tumour burden. Metastasised tumours were
 XX isolated from a site outside the renal capsule. RNA was isolated
 XX from primary tumours and metastases, reverse transcribed and
 XX subjected to differential display PCR. The sequences were analysed
 XX and compared with the sequences of the endogenous genes.
 XX present sequence. The method can be used to detect diagnosis and
 XX treat disorders related to metastasis, or treat malignant or
 XX non-malignant disorders, e.g. hyperplasia, dysplasia and
 XX hypertrophy. The metastatic sequence can be used to screen a

PR 16-NOV-1995; 95US-0006838.
 PA (THOM/) THOMPSON T.
 PI Thompson T;
 DX WPI: 1997-289397/26.
 XX Identifying tumour metastatic sequences - by introducing transfected
 PT cells into host mammal and analysing primary and metastatic
 PT sequences by differential display PCR
 XX Disclosure; Fig 12FN; 102pp; English.
 XX Mouse urogenital sinus (UGS) tissue was isolated from 17 day old
 CC mouse embryos. The UGS cells were infected with retroviruses,
 CC cultured and implanted under the renal capsule of mice.
 CC Reconstitutions were harvested 5 weeks later, when they showed
 CC signs of distress from the tumour burden. Metastasised tumours were
 CC from primary tumours outside the renal capsule. RNA was isolated
 CC from primary and metastasised tumours and subjected to differential
 CC display PCR. The sequences were analysed to obtain metastatic
 CC sequences, e.g. the sequence encoding the present sequence. The
 CC method can be used to detect, diagnose and treat disorders
 CC related to metastasis, e.g. hyperplasia, dysplasia and non-malignant
 CC disorders. The sequences can be used to screen a biologically
 CC sample for metastasis and the expression product can also be
 CC used to treat a metastatic disorder.
 XX Sequence 376 AA;
 SQ

Query Match 21.1%; Score 644.5; DB 18; Length 376;
 Identity 36.3%; Pident 40.1%; E 1e-101;
 Matches 213; Conservative 54; Mismatches 101; Indels 179; Gaps 47;

QY 16 GLPLAAKRFVDYGLHEDYFDRKRNQLRGWSDENEDSOLYPWRRGEGRWKDSWG 75
 DB 5 GAARDAAR-HDYGNRSAY--MRHNN--QMSDD-NDKNKY--YKRGDMRKMSKRG 53
 QY 76 GRVDAANTSGSPALCVSTFTVPLVPRCKEDAKNIVYERKSDLEASDPYVYK 135
 DB 54 GRV--AVTSDS--AVGSNT--AVN---RC--KDANGVY--KNCBN--ASSADYVYVW 96
 QY 136 TTGADDEDMENTSGOHLRFDPKPPPRPHGRKKNFYVVFHTLGOYFKGLQCSARVS 195
 DB 97 TAWSD---SGNGTSHHRY-DCK---HHGRMNHYVHYGYK-----GRCSVRVS 139
 QY 196 INTVLNVLGPOWMEVYPRHGRATPISKVDYVYITQDIPFYVYKQKDRSSDEYF 255
 DB 140 VNTANVT-----GMVTYVRHGRATV--AVKDVVVYVD---VVTM--KKNRSSDTK- 184
 QY 256 LDLPDFDFVLHDPSLPLNYSALYSKKNFQNDTGLEVSNHNTLNTVLTNCTFNFLYV 315
 DB 185 -----DHWAD-SR--NTST-NTKMS-DGNTG--YSTNITVNTVYV-NTGSR--IV 225
 QY 316 OTAVPGPCSPPTSPSSPSSPSPASSPSTLSTPSPSLMPTGYKSNLSIDNENCRIN 375
 DB 226 KAA--ACGRS-----KTSAGNDRDNC--NR 248
 QY 376 YGVFRATYITVDGLVLIOWADVPIPLTOPNSLMDFIYVTKGATPTACTATISDPFC 435
 DB 249 YGH-----ATTVG-----VWMDTWANSSD-----VVIC-----GSTVCTSDTC 282
 QY 436 QIAQNRVCSFVAVDELCLLVRRFNGSGTVCYVNTFLGDDASALTSISIPGRDLGSP 495
 DB 283 -----THTVCS-VVDV--NCTVRRP--NGSGTCYVN--TGDDTSATSTSV-----DRDAS-- 325
 QY 496 LRVYVGLISITGLAMFYVITLLIKKIKYTPKGNCRNVYVGGLSVFLSHAKAPPS 555
 DB 326 -RWANSA--SGVC-----NTVYSYKRYKHY-----NNSGNVYVRSKG---SVNRKAV-- 366
 QY 556 RGDREKD 562

DB 367 -GNKDN 372
 I:::I:
 RESULT 9
 AAY31978
 XX AAY31978 standard; Protein; 625 AA.
 AC AAY31978;
 XX 21-DEC-1999 (first entry)
 DT Mouse melanoma antigen gp100.
 XX Melanoma antigen gp100; mouse; antigen presentation;
 KW dendritic cell; adoptive immunotherapy; cancer; therapy; vaccine.
 OS Mus musculus.
 XX Key Location/Qualifiers
 FH Peptide 167-207
 FT Peptide /note= "minimal essential epitope"
 FT Peptide 208..216
 FT Peptide /note= "minimal essential epitope"
 FT Peptide 279..287
 FT Peptide /note= "minimal essential epitope"
 FT Peptide 422..425
 FT Peptide /note= "minimal essential epitope"
 FT Peptide 441..450
 FT Peptide /note= "minimal essential epitope"
 XX W09947103-A2.
 XX 23-SEP-1999.
 XX 19-MAR-1999; 99KO-US06031.
 XX 20-MAR-1998; 98US-0078890.
 XX (GENZ) GENZYME CORP.
 XX Nicolette CA, Kaplan J;
 XX WPI: 1995-590956/50.
 XX Preparing cells for use as cancer vaccines and in adoptive
 PT immunotherapy -
 XX Disclosure; Fig 2A; 55pp; English.
 CC The present sequence represents the mouse melanoma antigen gp100,
 CC a melanocyte differentiation antigen specifically recognised by
 CC HLA-A2 restricted tumour-infiltrating lymphocytes of melanoma
 CC patients. The invention provides methods for immunotherapy, in
 CC particular for inducing an immune response against an antigen in a
 CC patient. The invention also provides methods for inducing an immune
 CC response in a patient. The invention also provides methods for
 CC inducing an immune response in a patient. The invention also
 CC provides methods for inducing an immune response in a patient.
 CC are more potent presenters of exogenous peptide than the parental
 CC antigen-presenting cells are used. These APCs lack an effective
 CC endogenous TAP (transporter associated with antigen processing)
 CC activity and present exogenous antigen on the major
 CC histocompatibility complex class I (MHC-I) molecule. Suitable
 CC exogenous antigens include a tumour antigen, such as a minimal
 CC epitope, a tumour antigen, such as a minimal epitope, a
 CC generally modified APCs are useful for inducing an immune
 CC response (claimed) against an antigen in a patient (adoptive
 CC immunotherapy), especially as vaccines against cancer in mammals,
 CC preferably humans. The cells are also useful for expanding
 CC populations of immune effector cells, preferably cytotoxic T
 CC lymphocyte cells.
 XX Sequence 625 AA;
 SQ

Query Match 21.0%; Score 642; DB 20; Length 625;

```

Best Local Similarity 27.8%; Pred. No. 4e-50;
Matches 197; Conservative 89; Mismatches 191; Indels 178; Gaps 19;

Qy 54 MDQVLPVRRGSGHKWQCGQVQALGSDSPALVGSNTFTVNLVFPFCROKEDANGN 113
Db 43 MNRLVFPWTEVVG--SNCRWGQGVSLRVINDGPTLLIGANASFTSLRHPGSKVLPDQ 100
Qy 114 IYVENCRSOLELADSVYVNTKTTGADDEWEDNT---SQ---GQHRL---FP 157
Db 101 VI-----WANNIINGSOVMGQVYVPPQPDACVFP 132
Qy 158 DGKPPRPBHGKWKWVYVHTLCOVFOKLGCCSARVSNFTVNLVFGQVMEVTVFRBG 217
Db 133 DGGPCPSGPKPRSEVFNKMTGSKVQVGLGSPVSLRSIATGSHAKI.GTHNTEVTVHRRG 192
Qy 218 -RAYIPISKVDYVITDQIPFIVYQKNDNSDDETLRDLPIFDVLI.HDPSHFILNY 276
Db 230 SVYVPLHAKNSFTFTIDQVPSVSVSLOALQDGTAFHLNHLPLFALQGLHDPSGLAE 252
Qy 277 SALSIVKWSGNTFTIDQVPSVSVSLOALQDGTAFHLNHLPLFALQGLHDPSGLAE 252
Db 273 ADLSYTWDRGCGTGLTSLRALDTHVYLEGSVTAQVVLQAATLVLCVSGSPVPGTIDG 312
Qy 327 -----TPSPSSSTSPS-----PASSPSPILS----- 347
Db 313 MTFARAPGTTSGKGTTEWVGTTFGQMPFTTQSGTFTVGMPTFTVATISQMTSNVID 372
Qy 348 -----TPSPS-----LMPGYKSMELSDISNE----- 369
Db 373 TTLAEVSTTEGTTGTPTFTSTVVAQNTTTEGFDASPLLPSTOSGTGSPILLDDTDMILV 432
Qy 370 -----NCRINRGYFRATITVDGILEVNI IOAVDVPIPTLOPNSIMDFVTCKGATP 423
Db 433 KROYPLDCLVYRGFSALDITVOGIESABILQA--VPE---SEGDAFLNVSQGLP 486
Qy 424 TEACTIISDPTCOIAQRVCSFVAVDELCLLSVRRAP-NGSGTYCVNFTLGDASLALIS 482
Db 487 KEACMDIISPGCCPAORLCQVPSPPCQVILHOVLKGGSTGYCLNVLADANSLAVAS 546
Qy 483 ALISIPQKDLG---SELEHVNGVLISIGLAMPVMTVILLYKIKYKPGICNTRNVK 539
Db 547 TQLVVPQDGGGLQAPL-----LVGILLVAVVYASLTHRR-----LK 586
Qy 540 GKGLSVFLSHAKAPF-----SRGDREKDPIL 565
Db 587 KQGSVYSONPHGSHMLRIPVFRARGLGENSPIL 620

RESULT 10
AA78646
ID AA78646 standard; Protein: 661 AA.
AC
AA78646;
AA78646;
DE
22-JAN-1996 (first entry)
DE Melanoma associated antigen gp100.
KW
XX Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
XX Identification; tumour; gp100.
XX Homo sapiens.
XX EP668350-A1.
XX
PD 23-AUG-1995.
XX
XX 14-FEB-1995; 95EP-0200348.
XX
XX 21-DEC-1994; 94EP-0203709.
XX
XX 16-FEB-1994; 94EP-0200337.
XX
XX (ALKU) ARK2 NOBEL NV.

```

```

XX Adema GJ, Flygdr CG;
XX WPI: 1995-284750/38.
XX N-PSDB: AAQ96055.
XX Melanoma associated antigen gp100 - used in vaccines and for the
PT detection of tumours
XX Claim 1: Page 22-24; 40pp; English.
XX Immunogenic peptides derived from the melanoma associated antigen
CC encoding the immunogenic peptides may be used as primers and probes
CC in the detection of melanoma cells. Tumour infiltrating lymphocytes
CC capable of binding to the melanoma associated antigens can be
CC cultured ex vivo and returned to melanoma particles, and when
XX radiolabelled, they may be used to identify tumour deposits.
XX Sequence 661 AA;
Qy Query Match 20.5%; Score 627.5; DB 16; Length 661;
Best Local Similarity 27.1%; Pred. No. 9.5e-49;
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

Qy 7 VILVPLLAAGLPLQAAK--REFOVLEHQYQPDHRENNQLECHSDENEDQDQVYVRR 64
Db 9 LUHLAVICALLAVGATKVPKQMLG-----VSQRLR-----TKAMNRQLYPRWT- 53
Qy 65 GEGRMKNSDNGRGOAALTSDSPALVGSNTFTVNLVFPFCROKEDANGNVIYERN-CHSD 123
Db 54 -BAQRDCCMGGOVSLKVSNDGPTLLIGANASFTSLNFPGSKVLPDGOVIVWNTIING 112
Qy 124 LETASDPVYVNTKTTGADDEWEDNTSOGHLRPPDCKPFRPHGKWKWVYVHTLGOY 183
Db 113 SOVNGGQVYVPOET--DDA-----CIPFDGCGTSGMSKRSRFFVYVVKMTGQY 159
Qy 184 FQKLGCCSARVSNFTVNLVFGQVMEVTVFRBG-RAYIPISKVDYVITDQIPFIVTM 242
Db 160 MVLGQVPSVGLSGTIGRAMLGCTHTTEVTVHRRGSAFVPLANSSAFITDQVPSV 219
Qy 243 YOKNDNSDETFLRDLPIFDVLI.HDPSHFILNYSAISKWNGDNTGLEFSNNHILNHT 302
Db 220 SQLRALDGGKHFLRNOPLTFAQLHDPSGYLAEDLSYTMDFDSSGTLISRALVPH 279
Qy 303 VYLINGFENFI.VTQVAVP-----GPCP----- 324
Db 280 TLEPGVTVAGTVLQALPLTSCSSSPVGTTHGRTAPNTAGVPTTEVGTTHGQ 339
Qy 325 SPYPSPSST-----SPSPASSPSTLSTPSPSLM-----TG 357
Db 340 APTAEPSTGSVOVPTTEVISTAVPQMTAESTGCTPEKVPSEVMTTLAEMSTPEATG 399
Qy 358 YKSNELSDI-----SNF----- 369
Db 400 MTPAEVSVTLVSTAAQVTTTSMVETARELP.TPEPGFDASSIMSTESIGSLPLD 459
Qy 370 -----NCRINRGYFRATITVDGILEVNI IOAVDVPIPTLOPNSIMDFI 415
Db 460 GTATLRVLRQVDPDCLVYRGFSVFLDITVOGIESABILQA---VPSGSD--AFSLT 513
Qy 416 VTCKGTTEACTLIESDPTCOIAQRVCSFVAVDELCLLSVRRAP-NGSGTYCVNFTLGD 474
Db 514 VSCGGLPKEACMEISSPGCCPAORLCQVLPSPACQVILHOVLKGGSTGYCLNVLAD 573
Qy 475 DASIALTSALISIPQKDLG-SPURTVNGVLISIGLAMPVMTVILLYK 523
Db 574 TNSLAVVSTQLIMFGQAGLQVPLVIGIL-----LVMAVVLASLYRR 618

RESULT 11
AA731977
ID AA731977 standard; Protein: 661 AA.

```


[illegible]

DB		220	SQRLADGGKHKLEINOPLTALQHLDFSCYLAEADLSYTFMDPDSSTGLSLRALVVIHT	279
OY		303	VYLNGTFFNLIIVOTAP-----GPCP-----	324
DB		280	YLFCPTVAQVLQALHPIATSCGSSFWPGTGDCHRTAPARTAPNTAGQVTVEVGTTFQQ	339
OY		325	SPTSPSSST-----SPPASSPSPTLSTPSSLAP-----	357
DB		340	APTAEPDSGTISVQVPTTEVISTAPVQMTNESTGMTPEKVPVSEVMGTTLAEMSTPEATC	399
OY		358	KXMSLSDI-----	369
DB		400	MTPAESIVLSZTTAAQVITETWETTARELPTEPPSGDPASSINGESTIGSLPLLD	459
OY		370	-----NCIRNRGVFRATITVDGLEVINIQADVPIPTLOPNLSLMDFI	415
DB		460	GTAFLRKVRQLPVLCVLYRGSGFSLIDIVQGIESABILQA---VPSGBCD--AFELT	513
OY		416	VYCKGATPFACVILISOPTCIQNRVCSPVADEICLLSVRRAP-NGSGTYCVNFTLGD	474
DB		514	VSQOGLPKFAENELSSFCOPPAQLQCPVLRSPAQVHQIOLKGSGSYCLANSLAD	573
OY		475	DASLATLSALISIRKDLG-SPLRTNVGVILISICLAMFNVTWTILLYKK	523
DB		574	TNSLVAVSTQMIPQEGASLGVOPLVGIL----IYLMVAVLASLYR	618
DB		RESULT 14		
DB		AAU84803		
ID		AAU84803 standard; Protein: 561 AA.		
ACC		AAU84803;		
XX		08-MAY-2002 (first entry)		
XX		human gp100 consensus sequence.		
XX		savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;		
XX		viral infection; human immunodeficiency virus; melanoma;		
XX		bacterial infection; Salmonella; Legionella; parasitic infection;		
XX		Trypanosoma; Toxoplasma; Giardia.		
OS		Homo sapiens.		
XX		WO200150197-A1.		
XX		29-NOV-2001.		
PD		25-MAY-2001; 2001MO-UO0622.		
PF		26-MAY-2000; 2000AU-0007761.		
XX		(AUS) UNIV AUSTRALIAN NAT.		
PA		Thomson SA, Ramshaw IA;		
PPI		WPI; 2002-147575/19.		
XX		New synthetic polypeptides having several different segments of at		
PR		least one parent polypeptide linked together differently compared to		
PR		the said parent polypeptide, for inducing immune response		
XX		against a pathogen or cancer		
XX		Example 3; Fig 27; 364pp; English.		
XX		The invention relates to a new synthetic polypeptide (1) comprising		
CC		several different segments of at least one parent polypeptide linked		
CC		together in a different relationship relative to their linkage in the		
CC		parent polypeptide, for inducing immune response against a pathogen		
CC		function associated with the parent polypeptide and for inducing an		
CC		immune response against a pathogen or cancer. Also included are a		
CC		synthetic polynucleotide encoding and a computer system for		

Query Match	20.5%	Score 627.5	DB 23	Length 661
Heat/Local Similarity	27.1%	Pred: No. 9.5e-49		
Matches 176	Conservative 96	Mismatches 40	Indels 173	Gaps 19
7	VLAPELLAAGLPLOAK--RPRDVLGHEQYDPHNRNQLRSGDSDNDGGLYPWRR 64	1		
9	LHLAVLIGALLAVGATKVPKRMQGL-----VSRQLR-----TKANNRLYPEMT- 53	1		
65	GGRWKDSWEGGRVOAALTDSAPLGVSGNTFTVYVNIYPPRCOKEDANGVNIYERR-GRSD 123	1		
54	-DGRIDCWGGQVSLTGKADNDWDTLIGNARFSTGPGVQKLVGQVGVVYNTTLTG 112	1		
124	LFLASDPVYVYNTTGAUDDWEDNTSQGRLRPDPKPPRGRKKNKFNVPYPTLGYQ 183	1		
113	SOVWGGQVPYPOET--DDA-----CIPDGGCPGSGMSOKRSFVYKMTGWQY 159	1		
184	FOKLGCCGAKVGINVNTLVGPQWMEVYFRRHG-RAYPTISKVKQVYVITDQIPVPM 242	1		
160	MOYLGCPGSLGSLTGKMLCTHTVYVYHRRGKGVYLAHSSAFYITDOVPFVSQ 219	1		
243	YQKNDENSEDTFLRDLPTFFDVLVSHUFNYSALSYKWNFCNTGLGVSNHNLNLT 302	1		
220	SQLRALDGGKHKHFRNQPIITFALQHPDSCYLAELADLSYTMDFDSCSTLISRAUVYHT 279	1		
323	VYLVENFNKLTVQAVP-----GRCP----- 324	1		
280	TLRGQVTVQVQLNALPITSCGSSPVGCTGDRHTAEAWTTCQVPTTEWGTTCQ 339	1		
325	SPTSPFSST-----SPSASPSPSTLSPSLMP-----TG 357	1		
340	APTAEPSGTSVQVPTTEVISTAPVQMTAEOSTMTPEKVPVSEVMGTTLAENSTPATG 399	1		
358	YKSLUSD-----SNE----- 369	1		
400	MTPAEVSLVLSCTTAQVYTTBWTETAKELPIPPRPGDASSINSTESITCSGLPLD 459	1		
370	-----NKRINRYGVFRATITVDGLLEVALVLOADVPIPTLPQDNLMDPI 415	1		
460	GTAQLRLVKRQVPLDCLVYRGSFSTVLOIGTSEANLIQA-----VPSGEED--APELT 513	1		
414	VTCGATPTBACTYLLSDPTCOIAQNRVCSFPAVDELLCLSVRRFA-NGSGYCVWFTLGD 474	1		
516	VSCGGLFKKAEHLSPOCPQAPLQCVLPSPACQVLPSPACQVLPSPACQVLPSPACQV 573	1		
475	DASLATSALISPKDGLG-SPURTVNGYSIGCLAEFVWTVITLLYKK 523	1		
574	TNSIAVSVTQLLMQBQAGLQVPLVIGL-----LVTMAVYVASLIYRR 618	1		

Sequence: 661 AA:

Query Match: 20.5% Score 627.5 DB 23 Length 661

Heat/Local Similarity: 27.1% Pred: No. 9.5e-49

Matches 176 Conservative 96 Mismatches 40 Indels 173 Gaps 19

7 VLAPELLAAGLPLOAK--RPRDVLGHEQYDPHNRNQLRSGDSDNDGGLYPWRR 64

9 LHLAVLIGALLAVGATKVPKRMQGL-----VSRQLR-----TKANNRLYPEMT- 53

65 GGRWKDSWEGGRVOAALTDSAPLGVSGNTFTVYVNIYPPRCOKEDANGVNIYERR-GRSD 123

54 -DGRIDCWGGQVSLTGKADNDWEDNTSQGRLRPDPKPPRGRKKNKFNVPYPTLGYQ 112

124 LFLASDPVYVYNTTGAUDDWEDNTSQGRLRPDPKPPRGRKKNKFNVPYPTLGYQ 183

113 SOVWGGQVPYPOET--DDA-----CIPDGGCPGSGMSOKRSFVYKMTGWQY 159

184 FOKLGCCGAKVGINVNTLVGPQWMEVYFRRHG-RAYPTISKVKQVYVITDQIPVPM 242

160 MOYLGCPGSLGSLTGKMLCTHTVYVYHRRGKGVYLAHSSAFYITDOVPFVSQ 219

243 YQKNDENSEDTFLRDLPTFFDVLVSHUFNYSALSYKWNFCNTGLGVSNHNLNLT 302

220 SQLRALDGGKHKHFRNQPIITFALQHPDSCYLAELADLSYTMDFDSCSTLISRAUVYHT 279

323 VYLVENFNKLTVQAVP-----GRCP----- 324

280 TLRGQVTVQVQLNALPITSCGSSPVGCTGDRHTAEAWTTCQVPTTEWGTTCQ 339

325 SPTSPFSST-----SPSASPSPSTLSPSLMP-----TG 357

340 APTAEPSGTSVQVPTTEVISTAPVQMTAEOSTMTPEKVPVSEVMGTTLAENSTPATG 399

358 YKSLUSD-----SNE----- 369

400 MTPAEVSLVLSCTTAQVYTTBWTETAKELPIPPRPGDASSINSTESITCSGLPLD 459

370 -----NKRINRYGVFRATITVDGLLEVALVLOADVPIPTLPQDNLMDPI 415

460 GTAQLRLVKRQVPLDCLVYRGSFSTVLOIGTSEANLIQA-----VPSGEED--APELT 513

414 VTCGATPTBACTYLLSDPTCOIAQNRVCSFPAVDELLCLSVRRFA-NGSGYCVWFTLGD 474

516 VSCGGLFKKAEHLSPOCPQAPLQCVLPSPACQVLPSPACQVLPSPACQVLPSPACQV 573

475 DASLATSALISPKDGLG-SPURTVNGYSIGCLAEFVWTVITLLYKK 523

574 TNSIAVSVTQLLMQBQAGLQVPLVIGL-----LVTMAVYVASLIYRR 618

Result 15

AD ID AU009695

AD ID AU009695 standard; Protein: 661 AA.

AD ID AU009695;

Qy 243 YOKNRNDSDETFRLDLPFDVLIHDPHFNLNTSAISYKNNKNGDNTGLFVNNHNLNHT 302
Db 220 SQRALDGCNKHFLRQPLTFALQHDPSGLYLAEDSLTYWDRGDSGCTLSRSLVVTHT 279
Qy 303 YVLKNTFNFLVTQAVP-----GRCP----- 324
Db 280 YLEPGVTAAQVLOAAIPLTSCSSPVGCTTGDHRTAEAPNTTAGOVPTTEVGTGQ 339
Qy 325 SPTSPSSST-----SPSPASSPSTLSTSPSLMP-----TG 357
Db 340 APTAEPSGCTTSVOVPTTEVISTAPVQMPAEASTGMTPEKVPVSEVNGTTLAENSTPEATG 399
Qy 358 YKSNELSDI-----SNE----- 369
Db 400 MTPAEVSIVLGSCTTAQVTTENVTARELPPEPGPDASSIMSTESITGSLGPLD 459
Qy 370 -----NCRINRYGYFRATITVDGILEVNIQVADVPITPLQDPSLMDFI 415
Db 460 GTATLRLKROVPLDCLVLYRYGSFVTDIOVGIESAEILQA-----VPSGEGD--AFELT 513
Qy 416 YTKGATPEACTIISDPTQCOIAOHRVCSVPAVDEICLLSVRRAP--NGSGTYCVNFTGD 474
Db 514 VSCUGGLPKAEHESISGCGQPPAQRCLQVPLSPACQLVLIHQILKGGSGTYCLNVSLAD 573
Qy 475 DASLALTSALISIPGKDLG--SPLRTVNGVLISIGCLAMFVNTVITLLYKK 523
Db 574 TNSLAVYSTQLIMPQOAGLGQVPLVGIL-----LVLMAVYLASLIYRR 618

RESULT 7

US-09-267-439-121
; Sequence 121, Application US/09267439
; Patent No. 6270778
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; METHODS
; NUMBER OF SEQUENCES: 126
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L. L. P.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM: 3.5 INCH DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/267,439
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,174
; FILING DATE: 08/23/93
; APPLICATION NUMBER: US/08/231,565
; FILING DATE: 22-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4124US1
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELETYPE: 421722
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661
; TYPE: amino acid

; STRANDEDNESS: Unknown
; SOURCE: Unknown
; MOLECULE TYPE: Protein
US-09-267-439-121

Query Match 20.5%; Score 627.5; DB 4; Length 661;
Best Local Similarity 27.1%; Pred. No. 2.2e-52;

Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

Qy 7 VLVFLLAGLAPLOAAK--RFDVLIHGEQYPDHMRNNQLRQKSSDSEDEQIYVWRR 64
Db 9 LLHLAVTIGALLANGATKVPKPNQDLG-----VSQRLL-----TKAMRQRIPEWT- 53
Qy 65 GGGKWKSDMEGCKVQAAITSDSPALVGSMTFVVNIVPPRCOKEDANGNIVYERN--CRSD 123
Db 54 -EAPRLDCHVRGGVSKVSKNDGPTLIGANASFALNPFQSGKVLPGQGVWVNNITNG 112
Qy 124 ELASDPYVYVWMTGADDEWEDMTSOGUHLRFPDGKPFPHGRKKHVFYVHTLQY 183
Db 113 SQWNGGPVYQET--DIA-----CIPFDGQPCSSWSQSRSFYVWKTWQY 159
Qy 184 FQK--GGCSARYSNTVNTVYGVQVNEVYFRRHG--RAYIPISKVDVYITDIPVPTM 242
Db 160 MWV--GGFVSGISLGTGRAMLGTHTMEVTVYHRRGSRYSVPLAHSSAFTITDOVPESV 219
Qy 243 YOKNRNDSDETFRLDLPFDVLIHDPHFNLNTSAISYKNNKNGDNTGLFVNNHNLNHT 302
Db 220 SQRALDGCNKHFLRQPLTFALQHDPSGLYLAEDSLTYWDRGDSGCTLSRSLVVTHT 279
Qy 303 YVLKNTFNFLVTQAVP-----GRCP----- 324
Db 280 YLEPGVTAAQVLOAAIPLTSCSSPVGCTTGDHRTAEAPNTTAGOVPTTEVGTGQ 339
Qy 325 SPTSPSSST-----SPSPASSPSTLSTSPSLMP-----TG 357
Db 340 APTAEPSGCTTSVOVPTTEVISTAPVQMPAEASTGMTPEKVPVSEVNGTTLAENSTPEATG 399
Qy 358 YKSNELSDI-----SNE----- 369
Db 400 MTPAEVSIVLGSCTTAQVTTENVTARELPPEPGPDASSIMSTESITGSLGPLD 459
Qy 370 -----NCRINRYGYFRATITVDGILEVNIQVADVPITPLQDPSLMDFI 415
Db 460 GTATLRLKROVPLDCLVLYRYGSFVTDIOVGIESAEILQA-----VPSGEGD--AFELT 513
Qy 416 YTKGATPEACTIISDPTQCOIAOHRVCSVPAVDEICLLSVRRAP--NGSGTYCVNFTGD 474
Db 514 VSCUGGLPKAEHESISGCGQPPAQRCLQVPLSPACQLVLIHQILKGGSGTYCLNVSLAD 573
Qy 475 DASLALTSALISIPGKDLG--SPLRTVNGVLISIGCLAMFVNTVITLLYKK 523
Db 574 TNSLAVYSTQLIMPQOAGLGQVPLVGIL-----LVLMAVYLASLIYRR 618

RESULT 8

US-08-417-174-27
; Sequence 27, Application US/08417174
; Patent No. 5844075
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; METHODS
; NUMBER OF SEQUENCES: 126
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L. L. P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/417,174
 FILING DATE: 05-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/231,565
 FILING DATE: 22-APR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: CAROL M. GRUPPI
 REGISTRATION NUMBER: 37,341
 REFERENCE/DOCKET NUMBER: 2026-4124US1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 INFORMATION FOR SEQ. ID NO.: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 661
 TYPE: amino acid
 STRANDEDNESS: Unknown
 TOPOLOGY: Unknown
 MOLECULE TYPE: Protein

Query Match 20.5%; Score 626.5; DB 2; Length 661;

Best Local Similarity 27.1%; Pred. No. 2.8e-52;

Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

QY 7 VLVFLLAAGLPLQAAK--REFVDLGHIEQYPDHMRNNQLAGWSNDENWDQLYPVWR 64
 DB 9 LHLAVIGALLAVGATKVRNQWLGS-----VSQRLR-----TKANNQLIPMT- 53
 QY 65 GGRKKNSDEGGRVOALTSIDSPALVGSNTIYVNLVFPCKEDANGNVIYERN-CRSD 123
 DB 54 -EAQRLDCWRGGOVSLKVSNDGPTLIGANASFSIALNFGSKVLPDGOVIVNNTIING 112
 QY 124 LEASDPVYVNTTGAADDWEDNTSQGHLRFPDGPCKPPRPHGKKNVYVFTLGOY 183
 DB 113 SQWGGQVYVPOET--DDA-----CFFDGGPCPSGWSQKRSFYVWKTGQY 159
 QY 184 FQKLGQCSARVSIINTVNLVGPQWVIVFRHGG-RAYIPISKVKDYVYVITDQIPFTVM 242
 DB 160 WQFLGPGVSLGSICTGRAMLGTHTMEVTVYHRGSRVYVPLAHSSAFTITDQVPFSV 219
 QY 243 YQKNDRNSDETFLRDLPTFFDVLHDSHFHLYSAIYKKNFGDNTGLFVSNHNLNIT 302
 DB 220 SQLRALDGGKHFRLNQPTFAJLHDPFSGLAEADLSYINDGSGSTLISRALVYHT 279
 QY 303 YVLNGTFNLTQVATVP-----GFCP----- 324
 DB 280 YLEPGSVTAQVVLQAALPLTSCGSSPVGTCGHRPTAPANITACQVTEVVGTTGQ 339
 QY 325 SPTPSPSSST-----SPSPASSPSPTLSTPSPSLMP-----TG 357
 DB 340 APTNPSGTSVQVPTTEVISTAVQWPTAESTGHTPEKVPSEVWGTLTAEKSTPEATG 399
 QY 358 YKSNELSDI-----SNE----- 369
 DB 400 MTPAESVIVLSLSTTAAQVTTETWETARELPPEPEGPASSIMSTESITGSLGLD 459
 QY 370 -----NCRINRGYFRATITIVDGLVNIIOADVPIPTPLDNLMDFI 415
 DB 460 GTATRLKRVKQPLDVLIRGSGSVYLDIVGIESAEILR-----VPSGSD--AFETL 513
 QY 416 VTKCGATPCTACTIISDPTCOIAONRYCSPVAVDELCLLSVRFAF--NGSSTCYNFTLGD 474
 DB 514 VSCGGGLKEACMEISPECPQPAQLRCOPLPSPACQLVHLQLKGGSGTGYCLNLSVAD 573
 QY 475 DASIALTSALISIPKDLG--SPRLTVNGVLISIGCLAMEFVMTVITLLYKK 523

DB 574 TNSLAVSTQLINPQOEAGLGOVPLIVGIL-----LVLMVAVLASLIYER 618
 RESULT 9
 US-08-231-565A-27
 ; Sequence 27.5%; Pred. No. 2.8e-52;
 ; Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;
 ; GENERAL INFORMATION:
 ; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
 ; APPLICANT: STEVEN A.
 ; TITLE OF INVENTION: MELANOMA ANTIGENS AND
 ; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS: FINNEGAN
 ; STREET 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/231,565A
 ; FILING DATE: 22-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CAROL M. GRUPPI
 ; REGISTRATION NUMBER: 37,341
 ; REFERENCE/DOCKET NUMBER: 2026-4124
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; TELEX: 421792
 ; INFORMATION FOR SEQ. ID NO.: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 661
 ; TYPE: amino acid
 ; STRANDEDNESS: Unknown
 ; TOPOLOGY: Unknown
 ; MOLECULE TYPE: Protein
 ; US-08-231-565A-27

Query Match 20.5%; Score 626.5; DB 2; Length 661;

Best Local Similarity 27.1%; Pred. No. 2.8e-52;

Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

QY 7 VLVFLLAAGLPLQAAK--REFVDLGHIEQYPDHMRNNQLAGWSNDENWDQLYPVWR 64
 DB 9 LHLAVIGALLAVGATKVRNQWLGS-----VSQRLR-----TKANNQLIPMT- 53
 QY 65 GGRKKNSDEGGRVOALTSIDSPALVGSNTIYVNLVFPCKEDANGNVIYERN-CRSD 123
 DB 54 -EAQRLDCWRGGOVSLKVSNDGPTLIGANASFSIALNFGSKVLPDGOVIVNNTIING 112
 QY 124 LEASDPVYVNTTGAADDWEDNTSQGHLRFPDGPCKPPRPHGKKNVYVFTLGOY 183
 DB 113 SQWGGQVYVPOET--DDA-----CFFDGGPCPSGWSQKRSFYVWKTGQY 159
 QY 184 FQKLGQCSARVSIINTVNLVGPQWVIVFRHGG-RAYIPISKVKDYVYVITDQIPFTVM 242
 DB 160 WQFLGPGVSLGSICTGRAMLGTHTMEVTVYHRGSRVYVPLAHSSAFTITDQVPFSV 219
 QY 243 YQKNDRNSDETFLRDLPTFFDVLHDSHFHLYSAIYKKNFGDNTGLFVSNHNLNIT 302
 DB 220 SQLRALDGGKHFRLNQPTFAJLHDPFSGLAEADLSYINDGSGSTLISRALVYHT 279
 QY 303 YVLNGTFNLTQVATVP-----GFCP----- 324

Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 119;	
00Y	7 VLVFLTEAAGLPLQAAK--REFRDLVGEQYDDHRRNNQLQRKNSDENDEGLQVWRR 64
00Y	9 LLIILAILGALLAVGATKVPKIDMGLG-----VYSQLR-----TRANNRLQPYMT- 53
00Y	65 GEHRWDSGEGGQAALVSFSPALVGSNTFTFVNVLPYPPCKQDAGNVIVERN-CRSD 123
00Y	54 -ENRDLKRWGSGSLKVSNDQPTLIGANASFALNPQSQVILPQVKNVNTING 112
00Y	124 LEIASDPTVYNTTTCADDEDMNTSGOHLPLFPDCKFPFPRHCKKNVYVFHTLQY 163
00Y	113 SGLVGGQVPYQPT--DDA-----CIPFGDPCFSGWSQKSPVYKMTWGOY 159
00Y	184 FQHLGGCGARVSNVYLVGPQWVEVYFRRHG-NAVIPDKSVKQVYVITQDIPFVTH 242
00Y	160 WDFLPGVSGLSGTGRMLGLTHMTVYVYHRSQSVYVLAHSSAFTLIDVPPVS 219
00Y	243 YQKNDKNSDETFELRDLFPFDVLTHDPHFILYSAIYSKWFNGDGTGLFVSNMILNHT 302
00Y	220 SQLRALDGGKHFELRNQPLTEALQHDPSGLYASADLSYTFDFGDSGTLISRALVYTH 279
00Y	303 YVINGTFNPLVIOAVP-----GRCP----- 324
00Y	280 YLFGPYTAQVWLQALPLTSCSSPVPGTTHQHRPTAEAPNTAAQVPTTEVGTHPQ 339
00Y	325 SPTFSPSSST-----SPSPASPSPTLSPPSLMP-----TG 357
00Y	340 APTAEFGSTISGVQVPTTEVISTAPVOMPTAEISTGMPKRPVPSVGMKTILAEINSPETG 399
00Y	338 YKSMELSDI-----SNE----- 369
00Y	400 MTPAEVSIWLTAAQVTTTQWETTRARELPDPDGDASSIMSTESUSGLPLD 459
00Y	370 -----NCRINRGVFERATITVDGLLEVINIOTADVPIPLTLPQDNLMDFI 415
00Y	460 GTATLKVLRQVPLDCLVRYGVSFVSLDVLQVGLSASLIQA-----VPSGEGD--ATEFL 513
00Y	416 VTCGATPTACTTISDPTCIQAQRVCSFVAVDICLLSVRRFA-NSGSGYVCHVFGLD 474
00Y	514 VASGGLPCHMELSMSPGQCPQVPLPSPQAVLQHLQAGSLGCLVSLAD 573
00Y	475 DASLCTGAPKELSMIPQKDLG-SPLRTVYQVSLISIGLAEVNTVYLLYKK 523
00Y	574 TNSAVYVSTLIMPQAGLQGVPLVGLI-----LVLMVAVLASLYRR 618

RESULT 11
 US-09-267-439-27
 : Sequence 27, Application US/09267439
 : Patent No. 0270778
 :
 : GENERAL INFORMATION:
 : INVENTOR: KAWAMOTO, YUTAKA; ROSENBERG,
 : APPLICANT:
 : APPLICANT: STEVEN J.
 : TITLE OF INVENTION: MELANOMA ANTIGENS AND
 : TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
 : TITLE OF INVENTION: METHODS
 : NUMBER OF SEQUENCES: 126
 : COMPLETION DATE: 1998-08-11
 : ADDRESSEE: MORGAN, FINNEGAN, L.L.P.
 : STREET: 345 PARK AVENUE
 : CITY: NEW YORK
 : STATE: NEW YORK
 : COUNTRY: USA
 : ZIP CODE: 10022
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: FLOPPY DISK
 : COMPUTER: IBM PC COMPATIBLE
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: ASCII
 : CURRENT FILE: 09-267-439-27
 : APPLICATION NUMBER: US/09/267,439
 : FILING DATE:

Query Match	20.58;	Score 626.5;	DB 2;	Length 661;
Best Local Similarity	27.18;	Pred. No. 2.8e-52;		

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/417,174
;; FILING DATE: 05-APR-1995
;; PUBLICATION NUMBER: 08/08/231,565
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CAROL M. GRUPPI
;; REGISTRATION NUMBER: 37,341
;; REFERENCE/DOCKET NUMBER: 2026-412AUS1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 759-4800
;; TELEFAX: (212) 751-6849
;; INFORMATION FOR SEQ ID NO: 27:
;; LENGTH: 661
;; TYPE: amino acid
;; STRANDEDNESS: Unknown
;; TOPOLOGY: Unknown
;; MOLECULAR WEIGHT: Protein
us-09-267-439-27

Query Match 20.5%; Score 626.5; DB 4; Length 661;
Best Local Similarity 27.1%; Pred. No. 2,8e-52;
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

QY 7 VLVPLLAGLPLOAK--RPEDVLGHEOYFDHMRNNOLGSSDENHEDDOLYVWR 64
DB 9 LHLAVIGALLAVGATKVPKNDWLQ-----VSRQLR-----TKANRQLYPEWT- 53

QY 65 GEGRKDSREGKRVQALTSUSPALVGSNTFFVNLVFPCKEDANGVVERN-CHSD 133
DB 54 -ENQHLDCRGGVSLVSGDPLLGANASFIALNFFGSKVLPDQGVWNTING 112
QY 124 LELASDPYVNTWTTGADDEWEDTSGOHLARPDPGKPPRGRKKNFVVEHTLGOY 183
DB 113 SOVMGGQVPYQET--DDA-----CIFDGGCPGSCGSKSFVYVWKTGOY 159

QY 184 FOXLQCCSARVSTNVNLVGVQVMEVIVFRRIG-RAYIPISKVDVYVITDQIPVTM 242
DB 160 WQFLGVPVSLGIGTGKRALGTLCHTMVYTHRGSRVYFLNHSSTAFITDQVFSV 219

QY 243 YQKNDNRSDETLRLDIFDVLIDHPSHFLNYSALSYKWNFGDNTGLGVNNHPLNT 302
DB 220 SOLRALDGNKHEFLARQPLFTALQIDHPSGLAEADLSYTWDFGDSGSLTIRALVYTH 279

QY 303 VLVNTFENVTQVAP-----GCP- 324
DB 280 YLERGPVTAQVLAQAIPLSCGSSVPVGTDXDRPTAEPNTAQGVPTTEVGTITQ 339

QY 325 SPTSPSSST-----SPSSPSSPTLSTSPSSLP-----TG 357
DB 340 APTAEPSTGTVQVPTTEVISTAPVQMTAESTGMTPEKVPSEVYGTTLAENSTPEATG 399

QY 358 YKSMLESDI-----SNE 369
DB 400 MPAEVSIVLGSCTAAQVTTENVETARELPPEPEGPDASSIMSTESIGSLD 459

QY 370 -----NCRINRYGFRATITVIGLELVNIQVADVIFPTQLDNLMDPI 415
DB 460 GTATRLRVKROVPLDGVLYRGSEVTLDVIGIESAEILQA-----VPSGEGD-APELT 513

QY 416 VYCKATPTEACTISDPTQCIQANKRVSPVAVDELCLLSVRRAF-NSGTTCVNPTGLD 474
DB 514 VSCOGGLPKHEALMSEISSGCPQPAQQLVPSPAQQLVLIHLLKGSSTCLSLN 573

QY 475 DASLALTSALISTPKDGLG-SPLATVNGVLISIGGLAMFVMTVILLYKK 523
DB 574 TNSLVAVYVQLIMPQENGLGQVPLVIGL-----LVLMVAVLASIYRR 618

RESULT 12
us-07-891-9426-6

;; Sequence 6, Application US/078919426
;; Patent No. 5679511
;; GENERAL INFORMATION:
;; APPLICANT: Kwon, Young Se
;; TITLE OF INVENTION: KIM, KACK K. SPECIFIC GENE PMAL 17, MAPS
;; TITLE OF INVENTION: A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Christopher A. Michaels, Barnard, Brown &
;; STREET: 306 East State Street, Suite 220
;; CITY: Ithaca
;; STATE: NY USA
;; COUNTRY: USA
;; ZIP: 14850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/891,942G
;; FILING DATE: 01-JUN-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 06/0915,753
;; PUBLICATION NUMBER: 08/08/231,565
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/362,847
;; FILING DATE: 07-JUN-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Michaels, Christopher A
;; REGISTRATION NUMBER: 4,390
;; TELEPHONE: 607-273-1711
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 607-273-1711
;; TELEFAX: 607-273-2609
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 666 amino acids
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; PUBLICATION INFORMATION:
;; AUTHORS: Kwon, Young Se
;; AUTHORS: Chintamani, Chaya
;; AUTHORS: Chintamani, A
;; AUTHORS: Copeland, Neal G
;; AUTHORS: Gilbert, Debra J
;; AUTHORS: Jenkins, Nancy
;; AUTHORS: Barton, David
;; AUTHORS: Francke, Uta
;; AUTHORS: Kobayashi, Yvonne
;; AUTHORS: Kim, Kack K
;; TITLE: near the silver coat color locus on mouse
;; TITLE: chromosome 10 and is in a syntenic region on human
;; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
;; VOLUME: 88
;; PAGES: 9228-9232
;; DATE: October-1991
;; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 668
us-07-891-9426-6

Query Match 19.7%; Score 602; DB 1; Length 668;
Best Local Similarity 26.8%; Pred. No. 7e-50;
Matches 175; Conservative 95; Mismatches 212; Indels 170; Gaps 19;

QY 7 VLVPLLAGLPLOAK--RPEDVLGHEOYFDHMRNNOLGSSDENHEDDOLYVWR 64
DB 9 LHLAVIGALLAVGATKVPKNDWLQ-----VSRQLR-----TKANRQLYPEWT- 53

US-08-276-213-3

Query Match 4.3%, Score 131.5; DB 1; Length 521;
Local Similarity 20.9%, Frequent 00042;
Matches 63; Conservative 51; Mismatches 84; Indels 103; Gaps 13;
QY 230 YVTDQI-PIEVMTOKNRNSDFTFLRLDPTFDVLHDPHSHLYSAISYKWF--- 285
DB 269 YLFNONTAPWLGEGTITLSDTQTLMLKTL-----VQYLRLPT--AQYGADSEQMTFSM 321
QY 286 -----GNTGLFVNHNHLLHNTVLTGNTFNWTVQFAY-----PGPDSPTP 328
DB 322 NPDSGDTGILKDDQIVD--TVKDG--YLAPIKSSIFDPWASGASPSQSPFSYSHP 376
QY 329 SPSGSSPSPASPSPTLSTPSPLMPTFGYKGMELSDISNENCRINRYGFRATITVDG 388
DB 377 SPSFSASRTPPTPTPTAS--PIPTLPTA----- 404
QY 389 ILEVNIQIADVDPITLLOPNSLMDFIVTCKGAREACTIISDPTQAOAONVCSFYAV 448
DB 405 -----TPTPT-----ASPTPTTAAGCACTASYO-----VNS 432
QY 449 DELCLLSVRFAFGSG-----TYCVNFTLGDGASLALT--SALISIRCKDGLSGPLRTVNGV 502
DB 433 DWNGFTVTVAVINGSGVATKVTWTFNGTITNSNNAVTQNGSVTARNHYNV 492
QY 503 L 503
DB 493 I 493

RESULT 15

US-09-136-574A-43
US-09-136-5743 Application US/09136574A
Patent No. 6294366
GENERAL INFORMATION:
APPLICANT: Farrington, Graham K.
Anderson, Paige
Gibbs, Moreland
Berquist, Peter
Bantels, Roy
Williams, Hugh W.
Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing
Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS: Farrington, Graham K.
ANDERSON, PAIGE
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
SEQUENCE TYPE: Nucleotide
COMPUTER FILE: compatible
OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136.574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-5800

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1426 amino acids
STRAND: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6294366e
SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-136-574A-43

Query Match 4.0%, Score 123; DB 4; Length 1426;
Local Similarity 25.6%, Frequent 0.015;
Matches 44; Conservative 27; Mismatches 31; Indels 70; Gaps 10;
QY 315 VOTAVFGPCSP-----TPSPSSSTSPSPASSPSPTLS---TPSPSLMPTGYKSMEL 363
DB 952 IPTTPTPTTPTTVTVPTSPFPVSSSTPTTATPTPTPTTPTTPTTPTTPTTPTTPTT 1009
QY 364 SDTSNEN-----CRINRYGYFRATITVDGILEVH---IIQVAD- 399
DB 1010 TDDTNDMLFAGNKIVDKDKPKVLTVGNWEGENGT-NVPDGVNSCLKSALAEIANR 1068
QY 400 -----VPI-----PTLPQNSL--MDFIV-TCK 419
DB 1069 GFNLRFVPSAEILLNWSKGYKPKNINYYNPELGLTSLEVFDFVATCK 1120

Search completed: April 7, 2003, 16:12:44
Job time : 34 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:10:38 ; Search time 18 Seconds
(without alignments)

1842.767 Million cell updates/sec

Title: US-09-943-075A-2

Perfect score: 3061
Sequence: 1 MESICGVVLLANGLPQ.....PFSGKREKOPLODKFNNL 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published_Applications_AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US03_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US04_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2061	100.0	574	10	US-09-943-075A-2
2	2706	88.4	574	10	US-09-943-075A-2
3	2700	88.2	574	10	US-09-943-075A-2
4	2163	70.7	560	9	US-09-923-038A-36
5	2163	70.7	560	9	US-10-039-272-2
6	2163	70.7	560	10	US-09-735-705-225
7	2163	70.7	560	10	US-09-850-716A-225
8	2163	70.7	560	10	US-09-943-075A-25
9	2162	70.6	572	9	US-10-227-884-42
10	2162	70.6	572	9	US-10-230-163-42
11	2162	70.6	572	9	US-10-218-631-42
12	2162	70.6	572	9	US-10-230-338-42
13	2162	70.6	572	9	US-10-230-414-42
14	2162	70.6	572	9	US-09-812-238B-2
15	627.5	20.5	668	10	US-09-864-761-42194
16	620	20.3	626	9	US-10-047-539-2
17	462.5	15.1	105	10	US-09-864-761-34811
18	462.5	15.1	105	10	US-09-864-761-42194
19	295	9.6	69	10	US-09-864-761-42460

20	295	9.6	69	10	US-09-864-761-42460
21	208.5	5.8	98	10	US-09-925-301-1548
22	153	5.0	423	9	US-09-992-598-241
23	153	5.0	423	9	US-09-989-293A-241
24	153	5.0	423	9	US-09-989-735-241
25	153	5.0	423	9	US-09-990-444-241
26	153	5.0	423	9	US-09-989-730-241
27	153	5.0	423	9	US-09-990-444-241
28	133	5.0	423	9	US-09-991-181-241
29	133	5.0	423	9	US-09-993-687-241
30	133	5.0	423	9	US-09-989-734-241
31	133	5.0	423	9	US-09-997-653-241
32	133	5.0	423	9	US-10-174-590-180
33	133	5.0	423	9	US-10-175-758-180
34	133	5.0	423	9	US-09-993-667-243
35	133	5.0	423	9	US-10-173-706-180
36	153	5.0	423	9	US-10-175-738-180
37	153	5.0	423	9	US-10-175-752-180
38	153	5.0	423	9	US-10-176-482-180
39	153	5.0	423	9	US-10-176-511-180
40	153	5.0	423	9	US-10-180-553-180
41	153	5.0	423	9	US-10-180-557-180
42	153	5.0	423	9	US-10-180-557-180
43	153	5.0	423	9	US-09-990-438-241
44	153	5.0	423	9	US-09-990-562-241
45	153	5.0	423	9	US-09-990-562-241

ALIGNMENTS

RESULT 1
US-09-943-075A-2
Sequence 2, Application US/0943075A
; GENERAL INFORMATION:
; ORGANISM: US020151486A1
; APPLICANT: Safado, Payez F.
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Concomitlin Protein and Nucleic Acids Encoding the Same.
; FILE REFERENCE: 71369.282
; CURRENT APPLICATION NUMBER: US/09/943,075A
; PRIOR FILING DATE: 2001-08-30
; PRIORITY APPLICATION NUMBER: US 60/229,006
; SOFTWARE: SeqQuest
; SEQ ID NO 2
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Rat osteocalcin
US-09-943-075A-2

Query Match	100.0%	Score 3061	DB 10	Length 572
Best Local Similarity	100.0%	Pred. No. 5.2e-245	Mismatches 0	Indels 0
Matches 572	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	1	MESICGVVLLANGLPQQAARFVYLGHQYTPDHRENNOLGNSDENWDLQYP	60	
Db	1	MESICGVVLLANGLPQQAARFVYLGHQYTPDHRENNOLGNSDENWDLQYP	60	
OY	61	VYRGEGRKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPKCKEDAGNITVYRNC	120	
Db	61	VYRGEGRKDSWEGGRVQAALTSDSPALVGSNITFVNLVFPKCKEDAGNITVYRNC	120	
OY	121	RSLELASDPYVYNTTGADDEWENTSCQHLRFPDCKFPFRPHGRKKNFVYVHTL	180	
Db	121	RSLELASDPYVYNTTGADDEWENTSCQHLRFPDCKFPFRPHGRKKNFVYVHTL	180	
OY	181	GOYFQKLSQCARVSNITVNLVTGPQVNEVIFVRHGRAYTPIISKVDVIVITDQPIFV	240	

```
Db 181 GOYFOKLQGC SARVS INTVNLTVGPQMEVIVFRHGRAYIPISKVDVYVITDQIPFV 240
Qy 241 TMYOKNDNRMSDETFILRDLPIFDVLIHDPSSHFLANYSATSKYKNFQDNTGLFVSNHHTLN 300
Db 241 TMYOKNDNRMSDETFILRDLPIFDVLIHDPSSHFLANYSATSKYKNFQDNTGLFVSNHHTLN 300
Qy 301 HTYVLNGTFENLTVQTA VPGCPSPSPSSSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 360
Db 301 HTYVLNGTFENLTVQTA VPGCPSPSPSPSSSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 360
Qy 361 MEISDSINENCRINRYGYFRATITVDGILEVNIIOVADVPITLQPNLSLMDFIYCKAG 420
Db 361 MEISDSINENCRINRYGYFRATITVDGILEVNIIOVADVPITLQPNLSLMDFIYCKAG 420
Qy 421 ATPEACTIISDPTQICIAONRVCSPVAVDELCLLSVRRAFNGSGTVCYNFTLGDGASLAL 480
Db 421 ATPEACTIISDPTQICIAONRVCSPVAVDELCLLSVRRAFNGSGTVCYNFTLGDGASLAL 480
Qy 481 TSALISIPKDLGSPPLRYVNGVLISIGCLAVFMVTWITLLYKHKYKPKIGNCTRNRYNG 540
Db 481 TSALISIPKDLGSPPLRYVNGVLISIGCLAVFMVTWITLLYKHKYKPKIGNCTRNRYNG 540
Qy 541 KGLSVFLSHAKAPESRSGREKDPLODKPML 572
Db 541 KGLSVFLSHAKAPESRSGREKDPLODKPML 572

RESULT 2
US-09-943-075a-5
: Sequence 5, Application US/09943075a
: Best Local Similarity 89.04; Pred. No. 1.4e-215;
: Matches 508; Conservative 20; Mismatches 41; Indels 2; Gaps 1;
: GENERAL INFORMATION:
: APPLICANT: Popoff, Steven N.
: APPLICANT: Safado, Fayer F.
: APPLICANT: Smith, Steven L.
: APPLICANT: Snook, Steven L.
: TITLE OF INVENTION: Osteocalcin Protein and Nucleic Acids Encoding the Same,
: FILE REFERENCE: 71369.262
: CURRENT APPLICATION NUMBER: US/09/943, 075A
: PRIOR APPLICATION NUMBER: US/60/229, 006
: PRIOR FILING DATE: 2000-08-30
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 574
: ORGANISM: Mouse
US-09-943-075a-5
```

```
Db 241 TMSOKNDNRMSDETFILRDLPIFDVLIHDPSSHFLANYSATSKYKNFQDNTGLFVSNHHTLN 300
Qy 301 HTYVLNGTFENLTVQTA VPGCPSPSPSSSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 360
Db 301 HTYVLNGTFENLTVQTA VPGCPSPSPSPSSSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 360
Qy 359 KSMELSDISINENCRINRYGYFRATITVDGILEVNIIOVADVPITLQPNLSLMDFIYVC 418
Db 361 KSMELSDISINENCRINRYGYFRATITVDGILEVNIIOVADVPITLQPNLSLMDFIYVC 420
Qy 419 KGATPEACTIISDPTQICIAONRVCSPVAVDELCLLSVRRAFNGSGTVCYNFTLGDGASL 478
Db 421 KGATPEACTIISDPTQICIAONRVCSPVAVDELCLLSVRRAFNGSGTVCYNFTLGDGASL 480
Qy 479 ALTSALISIPKDLGSPPLRYVNGVLISIGCLAVFMVTWITLLYKHKYKPKIGNCTRNRY 538
Db 481 ALTSALISIPKDLGSPPLRYVNGVLISIGCLAVFMVTWITLLYKHKYKPKIGNCTRNRY 540
Qy 539 KKGGLSVFLSHAKAPESRSGREKDPLODKP 569
Db 541 KKGGLSVFLSHAKAPESRSGREKDPLODKP 571

RESULT 3
US-09-823-038A-36
: Sequence 36, Application US/09823038A
: Best Local Similarity 88.28; Pred. No. 3.8e-215;
: Matches 507; Conservative 20; Mismatches 42; Indels 2; Gaps 1;
: GENERAL INFORMATION:
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Abernethy, Nevil
: APPLICANT: Onrust, Rene
: APPLICANT: Kumble, Anand
: APPLICANT: Kumble, Anand
: APPLICANT: Kumble, Anand
: TITLE OF INVENTION: Compositions Isolated From Stromal Cells
: FILE REFERENCE: 11000.1037c3
: CURRENT APPLICATION NUMBER: US/09/823,038A
: CURRENT FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 36
: LENGTH: 574
: TYPE: PRT
: ORGANISM: Mouse
US-09-823-038A-36

Query Match 88.28; Score 2700; DB 10; Length 574;
Best Local Similarity 88.88; Pred. No. 3.8e-215;
Matches 507; Conservative 20; Mismatches 42; Indels 2; Gaps 1;

Qy 1 MESCGGVLFLLLAAGLPLOAKKRFVYLGHQYPDHMRNNOLRGWSSDENRDEHLYP 60
Db 1 MESCGGVLFLLLAAGLPLOAKKRFVYLGHQYPDHMRNNOLRGWSSDENRDEHLYP 60
Qy 61 VNRGEGWKSDEGGRVQALVSDSPALVGSNTTFVNLVFPFCQKEDANGNIYVERNC 120
Db 61 VNRGEGWKSDEGGRVQALVSDSPALVGSNTTFVNLVFPFCQKEDANGNIYVERNC 120
Qy 121 RSDLEASDPYVYVNTGADDEMDNTSQOHLRFPGDKPPRPHGRKKNFVYVFTL 180
Db 121 RNDLGLTSDLHYVNTAGADGDMEDTSSQHLRFDRPDRPFRPHGKKNFVYVFTL 180
Qy 181 GOYFOKLQGC SARVS INTVNLTVGPQMEVIVFRHGRAYIPISKVDVYVITDQIPFV 240
Db 181 GOYFOKLQGC SARVS INTVNLTVGPQMEVIVFRHGRAYIPISKVDVYVITDQIPFV 240
Qy 241 TMYOKNDNRMSDETFILRDLPIFDVLIHDPSSHFLANYSATSKYKNFQDNTGLFVSNHHTLN 300
Db 241 TMSOKNDNRMSDETFILRDLPIFDVLIHDPSSHFLANYSATSKYKNFQDNTGLFVSNHHTLN 300
Qy 301 HTYVLNGTFENLTVQTA VPGCPSPSPSSSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 360
```



```
Qy 421 ATPTACTIISDPTCQTAQNRVCSPVAVDELCLLSVRRFNGSGTCVNFITLGDASLAL 480
Db 407 SIPEVCTIISDPTCQTAQNRVCSPVAVDELCLLSVRRFNGSGTCVNFITLGDASLAL 466
Qy 481 TSALISPCGDLSPRTFNGVNLISGCLAMFVMTLLYKKHKYKPTGCTRWYVK 540
Db 467 TSTLISVPDRPASPLMANSALISVGCGLAIFVTVISLLVTKKHKYKPTGCTRWYVK 526

Qy 541 KGLSVFLSHAKAFKDRKDPILQDK 568
Db 527 KGLSVFLSHAKAFKDRKDPILQDK 554

RESULT 7
US-09-850-716a-225
: Sequence 225, Application US/09850716a
: Patent No. US2002011519a1
: GENERAL INFORMATION:
: APPLICANT: RSCG, Inc.
: APPLICANT: WarnerKris, Margalita
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedrick, Thomas S.
: APPLICANT: Carter, Patrick
: APPLICANT: Watanabe, Yoshihiro
: APPLICANT: Henderson, Robert A.
: APPLICANT: Peterson, David W.
: APPLICANT: Fanger, Neil
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.455C16
: CURRENT APPLICATION NUMBER: US/09/850,716a
: CURRENT FILING DATE: 2001-05-07
: NUMBER OF SEQ ID NOS: 440
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 225
: LENGTH: 560
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-850-716a-225

Query Watch 70.7%: Score 2163; DB 10; Length 560;
Best Local Similarity 69.4%; Pred. No. 9.8e-171;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

Qy 1 MESLCGVFLVLLAAGLPLQAKRFVGLGHEQYPOHRENNOLRWSSDENWDEQLYP 60
Db 1 MECLYFGLFLLAARLPIDANARFVIONERPYSATRENNOLRWSSDENWDEQLYP 60
Qy 61 VVRGEGRWKMEGGRVQAALTSQSPALVGSNITFVNLVFPCKEDANGNTVYERNK 120
Db 61 VVKRDMRWKMSKGRVQAVLTSDSPALVGSNITFAVNLFPCKEDANGNTVYERNK 120
Qy 121 RSLDLASDPVYNTGADDEWENTDSQOHLRFPOCKPTPHGKKNVYVFTL 180
Db 121 RNEAGLSADPVYNTANSEDSGNGTSGSHNVFPCKPTPHGKKNVYVFTL 180
Qy 181 GOYFQKLGCSARVSNITNLVTPGVMEVIFVRHGRAYIPISKVDVITDQIPFV 240
Db 181 GOYFQKLGCSARVSNITNLVTPGVMEVIFVRHGRAYIPISKVDVITDQIPFV 240
Qy 241 TWQKDRNSDPTLADLPITFDVLIHDSHFNLNSALSYKKNRGONTGLFVSNHTLN 300
Db 241 TWQKDRNSDPTLADLPITFDVLIHDSHFNLNSTYKKNRGONTGLFVSNHTLN 300
Qy 301 HTYVLNGTFENITVOTAVGPCSPPTSPSSSTSPSPASSPPTLSTPSLMPPTGYS 360
Db 301 HTYVLNGTFENITVOTAVGPCSPPTSPSSSTSPSPASSPPTLSTPSLMPPTGYS 360
Qy 361 MELSDINSEKRIKRYGPRATTIVGCLLEVNITQVADVITLQPDNSLMPVTCGK 420
Db 361 MELSDINSEKRIKRYGPRATTIVGCLLEVNITQVADVITLQPDNSLMPVTCGK 420
Qy 421 ATPTACTIISDPTCQTAQNRVCSPVAVDELCLLSVRRFNGSGTCVNFITLGDASLAL 480
Db 407 SIPEVCTIISDPTCQTAQNRVCSPVAVDELCLLSVRRFNGSGTCVNFITLGDASLAL 466
Qy 481 TSALISPCGDLSPRTFNGVNLISGCLAMFVMTLLYKKHKYKPTGCTRWYVK 540
Db 467 TSTLISVPDRPASPLMANSALISVGCGLAIFVTVISLLVTKKHKYKPTGCTRWYVK 526
```

```
Qy 541 KGLSVFLSHAKAFKDRKDPILQDK 568
Db 527 KGLSVFLSHAKAFKDRKDPILQDK 554

RESULT 7
US-09-897-778-225
: Sequence 225, Application US/09897778
: Patent No. US200201743A1
: GENERAL INFORMATION:
: APPLICANT: WarnerKris, Margalita
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedrick, Thomas S.
: APPLICANT: Carter, Patrick
: APPLICANT: Watanabe, Yoshihiro
: APPLICANT: Henderson, Robert A.
: APPLICANT: Peterson, David W.
: APPLICANT: Fanger, Neil
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.455C16
: CURRENT APPLICATION NUMBER: US/09/897,778
: CURRENT FILING DATE: 2001-06-26
: NUMBER OF SEQ ID NOS: 440
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 225
: LENGTH: 560
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-897-778-225

Query Watch 70.7%: Score 2163; DB 10; Length 560;
Best Local Similarity 69.4%; Pred. No. 9.8e-171;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

Qy 1 MESLCGVFLVLLAAGLPLQAKRFVGLGHEQYPOHRENNOLRWSSDENWDEQLYP 60
Db 1 MECLYFGLFLLAARLPIDANARFVIONERPYSATRENNOLRWSSDENWDEQLYP 60
Qy 61 VVRGEGRWKMEGGRVQAALTSQSPALVGSNITFVNLVFPCKEDANGNTVYERNK 120
Db 61 VVKRDMRWKMSKGRVQAVLTSDSPALVGSNITFAVNLFPCKEDANGNTVYERNK 120
Qy 121 RSLDLASDPVYNTGADDEWENTDSQOHLRFPOCKPTPHGKKNVYVFTL 180
Db 121 RNEAGLSADPVYNTANSEDSGNGTSGSHNVFPCKPTPHGKKNVYVFTL 180
Qy 181 GOYFQKLGCSARVSNITNLVTPGVMEVIFVRHGRAYIPISKVDVITDQIPFV 240
Db 181 GOYFQKLGCSARVSNITNLVTPGVMEVIFVRHGRAYIPISKVDVITDQIPFV 240
Qy 241 TWQKDRNSDPTLADLPITFDVLIHDSHFNLNSALSYKKNRGONTGLFVSNHTLN 300
Db 241 TWQKDRNSDPTLADLPITFDVLIHDSHFNLNSTYKKNRGONTGLFVSNHTLN 300
Qy 301 HTYVLNGTFENITVOTAVGPCSPPTSPSSSTSPSPASSPPTLSTPSLMPPTGYS 360
Db 301 HTYVLNGTFENITVOTAVGPCSPPTSPSSSTSPSPASSPPTLSTPSLMPPTGYS 360
Qy 361 MELSDINSEKRIKRYGPRATTIVGCLLEVNITQVADVITLQPDNSLMPVTCGK 420
Db 361 MELSDINSEKRIKRYGPRATTIVGCLLEVNITQVADVITLQPDNSLMPVTCGK 420
Qy 421 ATPTACTIISDPTCQTAQNRVCSPVAVDELCLLSVRRFNGSGTCVNFITLGDASLAL 480
Db 407 SIPEVCTIISDPTCQTAQNRVCSPVAVDELCLLSVRRFNGSGTCVNFITLGDASLAL 466
Qy 481 TSALISPCGDLSPRTFNGVNLISGCLAMFVMTLLYKKHKYKPTGCTRWYVK 540
Db 467 TSTLISVPDRPASPLMANSALISVGCGLAIFVTVISLLVTKKHKYKPTGCTRWYVK 526
```

QY 541 KGLSVFLSHAKAPESGDKREKOPLODK 568
Db 527 KGLSVFLNRAKAVFPGNQEKDLKLNQ 554

RESULT 8

US-09-943-075a-6
: Sequence 6, Application US/09943075A
: Patent NO. US20020151486A1
: GENERAL INFORMATION:
: APPLICANT: Desnoyers, Luc
: APPLICANT: Gerlitsen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gueney, Austin L.
: APPLICANT: St. Laurent, L.
: APPLICANT: Stephan, Jean-Philippe P.
: APPLICANT: Watanabe, Colin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: ACIDS ENCODING THE SAME
: CURRENT FILING DATE: 2002-08-26
: PRIOR FILING DATE: 1997-10-28
: NUMBER OF SEQ. ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 560
: TYPE: PRT
: ORGANISM: Human
US-09-943-075a-6

Query Match 70.7%; Score 2163; DB 10; Length 560;
Best Local Similarity 69.4%; Pred. No. 9,8e-171;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

QY 1 MESICGVVLVFLLAAGLPQANRFRVLGHEOYPDHRENNOLRGNSDENDEOLYP 60
Db 1 MESICGVVLVFLLAAGLPQANRFRVLGHEOYPDHRENNOLRGNSDENDEOLYP 60
QY 61 VHRGEGKRGKSGGGRVQAAVLTSSDPSLVGNSITFYVNYLPPKQCKEDANGNIVYERN 120
Db 61 VHRGEGKRGKSGGGRVQAAVLTSSDPSLVGNSITFYVNYLPPKQCKEDANGNIVYERN 120
QY 121 NSDLFLASDPYVYKATGADDEMDNFTSQGHLRFQKGRPRPRGKKNRIVYVHTL 180
Db 121 NSDLFLASDPYVYKATGADDEMDNFTSQGHLRFQKGRPRPRGKKNRIVYVHTL 180
QY 121 RNEAGLSADYVYVNTWNTANSESDSGENCTQSHNVFPGKPPHPQRRNRFIVFHTL 180
Db 121 RNEAGLSADYVYVNTWNTANSESDSGENCTQSHNVFPGKPPHPQRRNRFIVFHTL 180
QY 181 QYVFKLGQCSARNTVNLTVPGQVMEIVFRIGRAYIPIKVKDYVYVTDGPIFV 240
Db 181 QYVFKLGQCSARNTVNLTVPGQVMEIVFRIGRAYIPIKVKDYVYVTDGPIFV 240
QY 241 TWYQKNRNSDSEFRLDPIFFVLLVHDPDSFLYASISYKWFQDNTGLFVSNWTK 300
Db 241 TWYQKNRNSDSEFRLDPIFFVLLVHDPDSFLYASISYKWFQDNTGLFVSNWTK 300
QY 301 HTVYVLTGTFNHLVOTAVPQCPSPPTSPSSSTSPASSPSTLSTPSPSLMPTGYKS 360
Db 301 HTVYVLTGTFNHLVOTAVPQCPSPPTSPSSSTSPASSPSTLSTPSPSLMPTGYKS 360
QY 361 MELSDISNENCRINRGYFRATITVIGLLEVINIQVADVPITPLQPDNSLMDPFTCKG 420
Db 361 MELSDISNENCRINRGYFRATITVIGLLEVINIQVADVPITPLQPDNSLMDPFTCKG 420
QY 347 LESLRIDENCOINRGYFQATITVIGLLEVINIQVADVPITPLQPDNSLMDPFTCKG 406
Db 347 LESLRIDENCOINRGYFQATITVIGLLEVINIQVADVPITPLQPDNSLMDPFTCKG 406
QY 421 ATPTACTIISDPQIQANRQSPVAVDELCLLSVRRANGSGTCYVNTFLGDDASIAL 480
Db 421 ATPTACTIISDPQIQANRQSPVAVDELCLLSVRRANGSGTCYVNTFLGDDASIAL 480
QY 481 SATLSIPGKDLGSLPRTVNGVLIETGLAMFVYVITLLYKKIKYKPTGNCITRWVKG 540
Db 481 SATLSIPGKDLGSLPRTVNGVLIETGLAMFVYVITLLYKKIKYKPTGNCITRWVKG 540
QY 541 KGLSVFLSHAKAPESGDKREKOPLODK 568
Db 527 KGLSVFLNRAKAVFPGNQEKDLKLNQ 554

US-10-227-881-42
: Sequence 42, Application US/10227884
: Publication NO. US20030027988A1
: GENERAL INFORMATION:
: APPLICANT: Desnoyers, Luc
: APPLICANT: Gerlitsen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gueney, Austin L.
: APPLICANT: St. Laurent, L.
: APPLICANT: Stephan, Jean-Philippe P.
: APPLICANT: Watanabe, Colin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: ACIDS ENCODING THE SAME
: CURRENT FILING DATE: 2002-08-26
: PRIOR FILING DATE: 1997-10-28
: NUMBER OF SEQ. ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 42
: LENGTH: 560
: TYPE: PRT
: ORGANISM: Human
US-10-227-881-42

RESULT 9

US-10-227-881-42
: Sequence 42, Application US/10227884
: Publication NO. US20030027988A1
: GENERAL INFORMATION:
: APPLICANT: Desnoyers, Luc
: APPLICANT: Gerlitsen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gueney, Austin L.
: APPLICANT: St. Laurent, L.
: APPLICANT: Stephan, Jean-Philippe P.
: APPLICANT: Watanabe, Colin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: ACIDS ENCODING THE SAME
: CURRENT FILING DATE: 2002-08-26
: PRIOR FILING DATE: 1997-10-28
: NUMBER OF SEQ. ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 42
: LENGTH: 560
: TYPE: PRT
: ORGANISM: Human
US-10-227-881-42

```

; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095614
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100849
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106505
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12

```

```

; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

```

Query Match 70.6%; Score 2162; Db 9; Length 572;

Best Local Similarity 69.1%; Pred. No. 1.2e-170;

Matches 396; Conservative 78; Mismatches 87; Indels 12; Caps 2;

Qy 1 WSLGCVLFTLLAAGI PLQAAKRRFVGLGHDTFPHRRNNOLHGSSDKNKDEQLXP 60

Db 1 MECLYELGTLAARLP LIDAARKRFHVLGNRPSAYAREHNLGNSSDNDKMKLYP 60

Qy 61 VMRRGGKRWKDSMGKGVAAALTSDFALVGSNITFVNLVFPQCKEDANGNTVYERN 120

Db 61 VMRRGGKRWKDSMGKGVAAALTSDFALVGSNITFVNLVFPQCKEDANGNTVYERN 120

Qy 121 RSDLELASDPYVYVNTGTGADEWDENJTSOGHRLRPDCKPPRPHGRKKNNVYVFT 180

Db 121 RNEAGLSADPYVYVNTGTGADEWDENJTSOGHRLRPDCKPPRPHGRKKNNVYVFT 180

Oy 181 GOYFOKLGOCARVSINTVNLATVGPQWKEVIVFRHGRATIPISKVKVDVITDQIPFV 240
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
Db 181 GOYFOKLGOCARVSINTVNLATVGPQWKEVIVFRHGRATIPISKVKVDVITDQIPFV 240
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
Oy 241 TAYKONDRSSDTEFLKOLPFDVLIHDSHETLKYSAIKNFGDWTGLYVSNHTLN 300
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
Db 241 TAYKONDRSSDTEFLKOLPFDVLIHDSHETLKYSAIKNFGDWTGLYVSNHTLN 300
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-17
Oy 301 HTVVLNGTNEFNITVQTAVGCGESPTSPSSSTSPASSPSTLST-----PSFLMP 355
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
Db 301 HTVVLNGTNEFNITVQTAVGCGESPTSPSSSTSPASSPSTLST-----PSFLMP 355
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
Oy 355 TGKWSMELLSIDSNENRINRYGYFRATITVCGILLENLIVQADVPITLOPDSNWDFT 415
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
Db 355 TGKWSMELLSIDSNENRINRYGYFRATITVCGILLENLIVQADVPITLOPDSNWDFT 415
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
Oy 415 VTCKATPTPEACHTISDPTCOATONRVCSVPVADCLLSVRFRFNGSGTYCVNFTLGD 475
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-18
Db 415 VTCKATPTPEACHTISDPTCOATONRVCSVPVADCLLSVRFRFNGSGTYCVNFTLGD 475
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-18
Oy 475 ASLATSALISITKGDGLSPKTVNGVLISITGLAMEVMTVITLLYKHKHTVKGTCNTR 535
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
Db 475 ASLATSALISITKGDGLSPKTVNGVLISITGLAMEVMTVITLLYKHKHTVKGTCNTR 535
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
Oy 535 NVVREKGLSVFISHAKAPFRCORENDREXPLLODK 568
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
Db 535 NVVREKGLSVFISHAKAPFRCORENDREXPLLODK 568
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
RESULT 10
us-10-230-163-42
? Sequence 42, Application US/10230163
? GENERAL INFORMATION
? FILE REFERENCE: P35502199
? CURRENT APPLICATION NUMBER: US/10/230,163
? CURRENT FILING DATE: 2002-08-28
? PRIOR FILING DATE: 2002-04-09
? PRIOR APPLICATION NUMBER: 60/059113
? PRIOR FILING DATE: 1997-10-17
? PRIOR APPLICATION NUMBER: 60/062287
? PRIOR FILING DATE: 1997-10-17
? PRIOR APPLICATION NUMBER: 60/063549
? PRIOR FILING DATE: 1997-10-28
? PRIOR APPLICATION NUMBER: 60/064103
? PRIOR FILING DATE: 1997-10-31
? PRIOR APPLICATION NUMBER: 60/069873
? PRIOR FILING DATE: 1998-03-20
? PRIOR APPLICATION NUMBER: 60/078910
? PRIOR FILING DATE: 1998-03-20
? PRIOR APPLICATION NUMBER: 60/079294
? PRIOR FILING DATE: 1998-03-25
? PRIOR APPLICATION NUMBER: 60/079656
? PRIOR FILING DATE: 1998-03-26
? PRIOR APPLICATION NUMBER: 60/079728
? PRIOR FILING DATE: 1998-03-27
? PRIOR APPLICATION NUMBER: 60/081819
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/081955
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/082804
? PRIOR FILING DATE: 1998-04-22
? PRIOR APPLICATION NUMBER: 60/084441
? PRIOR FILING DATE: 1998-05-06
? PRIOR APPLICATION NUMBER: 60/085323
? PRIOR FILING DATE: 1998-05-17
? PRIOR APPLICATION NUMBER: 60/085579
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/086392
? PRIOR FILING DATE: 1998-05-22
? PRIOR APPLICATION NUMBER: 60/089538
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089538
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089538
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/090472
? PRIOR FILING DATE: 1998-06-18
? PRIOR APPLICATION NUMBER: 60/090472
? PRIOR FILING DATE: 1998-06-18
? PRIOR APPLICATION NUMBER: 60/090691
? PRIOR FILING DATE: 1998-06-25
? PRIOR APPLICATION NUMBER: 60/090691
? PRIOR FILING DATE: 1998-06-25
? PRIOR APPLICATION NUMBER: 60/090695
? PRIOR FILING DATE: 1998-06-25
? PRIOR APPLICATION NUMBER: 60/090695
? PRIOR FILING DATE: 1998-06-25
? PRIOR APPLICATION NUMBER: 60/091982
? PRIOR FILING DATE: 1998-07-07
? PRIOR APPLICATION NUMBER: 60/091982
? PRIOR FILING DATE: 1998-07-07
? PRIOR APPLICATION NUMBER: 60/095302
? PRIOR FILING DATE: 1998-08-04
? PRIOR APPLICATION NUMBER: 60/095302
? PRIOR FILING DATE: 1998-08-04
? PRIOR APPLICATION NUMBER: 60/095318
? PRIOR FILING DATE: 1998-08-04
? PRIOR APPLICATION NUMBER: 60/095318
? PRIOR FILING DATE: 1998-08-04
? PRIOR APPLICATION NUMBER: 60/095916
? PRIOR FILING DATE: 1998-08-10
? PRIOR APPLICATION NUMBER: 60/095916
? PRIOR FILING DATE: 1998-08-10
? PRIOR APPLICATION NUMBER: 60/096146
? PRIOR FILING DATE: 1998-08-17
? PRIOR APPLICATION NUMBER: 60/096146
? PRIOR FILING DATE: 1998-08-17
? PRIOR APPLICATION NUMBER: 60/096791
? PRIOR FILING DATE: 1998-08-17
? PRIOR APPLICATION NUMBER: 60/096791
? PRIOR FILING DATE: 1998-08-17
? PRIOR APPLICATION NUMBER: 60/097586
? PRIOR FILING DATE: 1998-08-26
? PRIOR APPLICATION NUMBER: 60/097586
? PRIOR FILING DATE: 1998-08-26
? PRIOR APPLICATION NUMBER: 60/098544
? PRIOR FILING DATE: 1998-09-31
? PRIOR APPLICATION NUMBER: 60/098544
? PRIOR FILING DATE: 1998-09-31
? PRIOR APPLICATION NUMBER: 60/099596
? PRIOR FILING DATE: 1998-09-09
? PRIOR APPLICATION NUMBER: 60/099596
? PRIOR FILING DATE: 1998-09-09
? PRIOR APPLICATION NUMBER: 60/099598
? PRIOR FILING DATE: 1998-09-09
? PRIOR APPLICATION NUMBER: 60/099598
? PRIOR FILING DATE: 1998-09-09
? PRIOR APPLICATION NUMBER: 60/099803
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099803
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099811
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099811
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099812
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099812
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099816
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/099816
? PRIOR FILING DATE: 1998-09-10
? PRIOR APPLICATION NUMBER: 60/100038
? PRIOR FILING DATE: 1998-09-11
? PRIOR APPLICATION NUMBER: 60/100038
? PRIOR FILING DATE: 1998-09-11
? PRIOR APPLICATION NUMBER: 60/100385
? PRIOR FILING DATE: 1998-09-15
? PRIOR APPLICATION NUMBER: 60/100385
? PRIOR FILING DATE: 1998-09-15
? PRIOR APPLICATION NUMBER: 60/100390
? PRIOR FILING DATE: 1998-09-15
? PRIOR APPLICATION NUMBER: 60/100390
? PRIOR FILING DATE: 1998-09-15
? PRIOR APPLICATION NUMBER: 60/100627
? PRIOR FILING DATE: 1998-09-16
? PRIOR APPLICATION NUMBER: 60/100627
? PRIOR FILING DATE: 1998-09-16
? PRIOR APPLICATION NUMBER: 60/100848
? PRIOR FILING DATE: 1998-09-17
? PRIOR APPLICATION NUMBER: 60/100848
? PRIOR FILING DATE: 1998-09-17
? PRIOR APPLICATION NUMBER: 60/100919
? PRIOR FILING DATE: 1998-09-17
? PRIOR APPLICATION NUMBER: 60/100919
? PRIOR FILING DATE: 1998-09-17
? PRIOR APPLICATION NUMBER: 60/101477
? PRIOR FILING DATE: 1998-09-23
? PRIOR APPLICATION NUMBER: 60/101477
? PRIOR FILING DATE: 1998-09-23
? PRIOR APPLICATION NUMBER: 60/101738
? PRIOR FILING DATE: 1998-09-24
? PRIOR APPLICATION NUMBER: 60/101738
? PRIOR FILING DATE: 1998-09-24
? PRIOR APPLICATION NUMBER: 60/101741
? PRIOR FILING DATE: 1998-09-27
? PRIOR APPLICATION NUMBER: 60/101741
? PRIOR FILING DATE: 1998-09-27
? PRIOR APPLICATION NUMBER: 60/101786
? PRIOR FILING DATE: 1998-04-15

		Query Match	Similarity	70.5%	Score 2162	DB 9	Length 572
		Matches	396	Conservative	78	Mismatches 87	Indels 12
						Gaps	
1	PRIOR APPLICATION NUMBER: 60/149638						
2	PRIOR FILING DATE: 1999-08-17						
3	PRIOR APPLICATION NUMBER: 60/151733						
4	PRIOR FILING DATE: 1999-08-31						
5	PRIOR APPLICATION NUMBER: 60/164418						
6	PRIOR FILING DATE: 1999-11-09						
7	PRIOR APPLICATION NUMBER: 60/166361						
8	PRIOR FILING DATE: 1999-11-09						
9	PRIOR APPLICATION NUMBER: 60/169445						
10	PRIOR FILING DATE: 1999-12-07						
11	PRIOR APPLICATION NUMBER: 60/169495						
12	PRIOR FILING DATE: 1999-12-07						
13	PRIOR APPLICATION NUMBER: 60/169835						
1	1 MESSAGVVLVLLLAAGLPLAAAKAFRDVLIGHEQYPUHNRNQLRQSGSDENMDQLXP 60						
2	1 MECLYVFLAGLLLAARLPDLAAKAFHDVLIGHEPQYATYMRHQLNCGSDENMDKLYP 60						
3	1 WNRGKQKDSHGRGQVQAULTSDSPALVGSNTFYVNLVYPRCKQKEDANGNIVYERKNC 120						
4	1 VVKHEDHAKNMKGKGVQVAULTSDSPALVGSNTFYVNLVYPRCKQKEDANGNIVYERKNC 120						
5	121 RSDLELASDPVYVNMVTCADDEDMDSTSQOQLRFPDGRPBPGRKKNNVYVPHL 180						
6	121 RNFAGLSAOPVYVNMVTCADDEDMDSTGQSHNVFPDGRPBPGRKKNNVYVPHL 180						
7	181 GYFQKLGRCVSRYVNTANTYLPGLVEVYVTRBGRAYVZIAQVKDYVVTQIPVY 240						
8	181 GYFQKLGRCVSRYVNTANTYLPGLVEVYVTRBGRAYVZIAQVKDYVVTQIPVY 240						
9	241 TMYQKNDKNSDFTLRQLPFEFVLTHDPDSHFLANYSALSKNFGDNTGLFVSNHNL 300						
10	241 TMYQKNDKNSDFTLRQLPFEFVLTHDPDSHFLANYSALSKNFGDNTGLFVSNHNL 300						
11	301 HTYVLNKTSLNLTVKAAGDPCPPPPPPR-----PSKPTSLTSLTSDSNTPGP 355						
12	356 TGYKSMELSDSNENCRNRYGYRATITIVDGLSEVNIQVADPIPTLPQDNSLPGI 415						
13	356 TGYKSMELSDSNENCRNRYGYRATITIVDGLSEVNIQVADPIPTLPQDNSLPGI 415						
14	413 TGNPLSLRGPIDENCOINRGYQATITIVBGLSEVNIQVADPIPTLPQDNSLPGI 473						
15	413 TGNPLSLRGPIDENCOINRGYQATITIVBGLSEVNIQVADPIPTLPQDNSLPGI 473						
16	416 VTCRKAPTEACTIISDPTQIAQNRVSVYVNDGLCLSVHRAFGSGYQVNTLGGD 475						
17	416 VTCRKAPTEACTIISDPTQIAQNRVSVYVNDGLCLSVHRAFGSGYQVNTLGGD 475						
18	414 VTCQSGTPEVNTCTIISDPTQIAQNRVSVYVNDGLCLSVHRAFGSGYQVNTLGGD 473						
19	476 ASLATSALISIPKQGLSPRTYVNGVLISGLAMVYVNTVILLYKKHKTYPGTGMR 535						
20	476 ASLATSALISIPKQGLSPRTYVNGVLISGLAMVYVNTVILLYKKHKTYPGTGMR 535						
21	474 TSLSATLSLISYDRDPASPRLMANSALISVGCGLFAIVTSLVLTKKHKNYIENS 533						
22	474 TSLSATLSLISYDRDPASPRLMANSALISVGCGLFAIVTSLVLTKKHKNYIENS 533						
23	536 NVYKQKGLSVFLSHAHAFSGRGRKQPLQDX 568						
24	536 NVYKQKGLSVFLSHAHAFSGRGRKQPLQDX 568						
25	534 NVYKQKGLSVFLSHAHAFSGRGRKQPLQDX 566						
26	534 NVYKQKGLSVFLSHAHAFSGRGRKQPLQDX 566						
1	RESULT 11						
2	US-10-218-631-42						
3	Sequence 42, Application US/10218631						
4	Publication No. US20030045667A1						
5	Publication Date: 2003-02-27						
6	APPLICANT: Baker, Kevin P.						
7	APPLICANT: Desnoyers, Luc						
8	APPLICANT: Gerltsen, Mary						
9	APPLICANT: Goddard, Audrey						
10	APPLICANT: Godowski, Paul J.						
11	APPLICANT: Grimaldi, J. Christopher						
12	APPLICANT: Hare, John L.						
13	APPLICANT: Harvey, Victoria						
14	APPLICANT: Steglin, Jean-Philippe F.						

APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
CURRENT APPLICATION NUMBER: US/10/218,631
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-26
Remaining of SEQ ID NOS: 246
Remainder of SEQ ID NOS: 246
SEQ ID NO 42
LENGTH: 572
TYPE: PRT
ORGANISM: Homo Sapien
US-10-218-631.42

Query Match 70.6%; Score 2162; DB 9; Length 572;
Best Local Similarity 69.1%; Pred. No. 1.2e-170; Indels 12; Gaps 2;
Matches 396; Conservative 78; Mismatches 87;

Qy 1 MESLGGVYVLLAAGLPQAKRFVDLGHBOYVDHRENNOLRGSSDENMDQLYP 60
Db 1 MCLYVFLGFLLAARLPDAARFHDVLGNRPVAYRHHNOLRGSSDENMDKSLYP 60
Qy 61 VNRGEGRWKDSMGEGRVQAALSDSPALVGSNITFVNLVFPKCKEDANGNTVYERNK 120
Db 61 VNRGEGRWKDSMGEGRVQAALSDSPALVGSNITFVNLVFPKCKEDANGNTVYERNK 120
Qy 121 RSDLEASDPVYVNTMTGDEEDNTSGOGRHFRPCKGPPRPHGRKKNVYVHTL 180
Db 121 RNEAGLSADPVYVNTMTGDEEDNTSGOGRHFRPCKGPPRPHGRKKNVYVHTL 180
Qy 181 GQYFQKLGCSARVSTINTVNLVGPQWMEIVFRHGRAYITPSKYKQVYVTDQIPFV 240
Db 181 GQYFQKLGCSARVSTINTVNLVGPQWMEIVFRHGRAYITPSKYKQVYVTDQIPFV 240
Qy 241 TWYKNDNRSSDETFRLDQIFDVLHDPSEFLAYSNKYKNGFQNTGLVPSNHTL 300
Db 241 TWYKNDNRSSDETFRLDQIFDVLHDPSEFLAYSNKYKNGFQNTGLVPSNHTL 300
Qy 301 HTVYLVNGTINFLVOTAVGCPSTPTSPSSSTSPASSPTLST ----PSPSLMP 355
Db 301 HTVYLVNGTINFLVOTAVGCPSTPTSPSSSTSPASSPTLST ----PSPSLMP 355
Qy 356 TGYKSMWELSDINENCRNYGFRATITVDCGILEVNIIOADVPITLQPNLSMDPT 415
Db 356 TGYKSMWELSDINENCRNYGFRATITVDCGILEVNIIOADVPITLQPNLSMDPT 415
Qy 354 TGDNPFLSRIPDNCQINRYGHQMTITIVGCIENVN11QMTDVLMPVPFESSLIDFV 413
Db 354 TGDNPFLSRIPDNCQINRYGHQMTITIVGCIENVN11QMTDVLMPVPFESSLIDFV 413
Qy 416 VTCGATPTACTITSDPTQCAQNRVCSVPVADCLGLSVRAFNGSCTVCYNFTIGDD 475
Db 416 VTCGATPTACTITSDPTQCAQNRVCSVPVADCLGLSVRAFNGSCTVCYNFTIGDD 475
Qy 414 VTCGGSIPFVCTIISDPTCEITQNTVCSQVDDMCLLTVHRTFNGSGTVCNLTIGDD 473
Db 414 VTCGGSIPFVCTIISDPTCEITQNTVCSQVDDMCLLTVHRTFNGSGTVCNLTIGDD 473
Qy 476 ASLATSALISTGCKDLSRLTNGVLTISTGCLAMVMTVTLAYKKHTKPTGCTR 535
Db 476 ASLATSALISTGCKDLSRLTNGVLTISTGCLAMVMTVTLAYKKHTKPTGCTR 535

Db 474 TSLATSTLISVDPDRPASPLRMANSALISVGLAIFVTIVTISLLVYKKHKEVNFISNPG 533
Qy 536 NVYKGLGSVFLSHAKAPFSRGDRKDPDLQOK 568
Db 534 NVYKGLGSVFLNRKAKAVFFPGDKRDLPLKNQ 566
RESULT 12
US-10-230-338-42
Sequence 42, Application US/10203038
Publication No. US20030044934A1
GENERAL INFORMATION:
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
APPLICANT: Berncoyars, Marc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Slepchak, Victor
APPLICANT: Slepchak, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P350P1C92
CURRENT APPLICATION NUMBER: US/10/230,338
CURRENT FILING DATE: 2002-08-28
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-12-17/078910
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17/078910
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-26
Remaining of SEQ ID NOS: 246
Remainder of SEQ ID NOS: 246
SEQ ID NO 42
LENGTH: 572
TYPE: PRT
ORGANISM: Homo Sapien
US-10-230-338-42

Query Match 70.6%; Score 2162; DB 9; Length 572;
Best Local Similarity 69.1%; Pred. No. 1.2e-170; Indels 12; Gaps 2;
Matches 396; Conservative 78; Mismatches 87;

Qy 1 MESLGGVYVLLAAGLPQAKRFVDLGHBOYVDHRENNOLRGSSDENMDQLYP 60
Db 1 MCLYVFLGFLLAARLPDAARFHDVLGNRPVAYRHHNOLRGSSDENMDKSLYP 60
Qy 61 VNRGEGRWKDSMGEGRVQAALSDSPALVGSNITFVNLVFPKCKEDANGNTVYERNK 120
Db 61 VNRGEGRWKDSMGEGRVQAALSDSPALVGSNITFVNLVFPKCKEDANGNTVYERNK 120
Qy 121 RSDLEASDPVYVNTMTGDEEDNTSGOGRHFRPCKGPPRPHGRKKNVYVHTL 180
Db 121 RNEAGLSADPVYVNTMTGDEEDNTSGOGRHFRPCKGPPRPHGRKKNVYVHTL 180
Qy 181 GQYFQKLGCSARVSTINTVNLVGPQWMEIVFRHGRAYITPSKYKQVYVTDQIPFV 240
Db 181 GQYFQKLGCSARVSTINTVNLVGPQWMEIVFRHGRAYITPSKYKQVYVTDQIPFV 240

181	QY	GOYFOGLGRCRVYRVYNTANVTLGOLMEVTVYRREGRVAVPAQGVQVYVVDIOLPVEV	240
241	QY	TMVKQNRNSDETFELROLPIFFDVLHDPSPHPLYSALSYKNNFGDGTCLGVSNHNTLN	300
241	QY	YKNNHNTLNHNTLNHNTLNHNTLNHNTLNHNTLNHNTLNHNTLNHNTLNHNTLNHNTLN	300
241	QY	THQNRNSDETFELROLPIFFDVLHDPSPHPLYSALSYKNNFGDGTCLGVSNHNTLN	300
301	QY	HTVYKNTKNTNLVAVGVCSPSTFSPSSGSPASPSFTLSL-----PSPSLMP	355
301	QY	HTVYKNTKNTNLVAVGVCSPSTFSPSSGSPASPSFTLSL-----PSPSLMP	355
301	QY	HTVYKNTKNTNLVAVGVCSPSTFSPSSGSPASPSFTLSL-----PSPSLMP	355
356	QY	TCYKNGELSDSNMCRNCKYCFRATITVCGELVENVLIQVADYPTIOLDSNLMQETI	413
356	QY	TCYKNGELSDSNMCRNCKYCFRATITVCGELVENVLIQVADYPTIOLDSNLMQETI	413
356	QY	TCYKNGELSDSNMCRNCKYCFRATITVCGELVENVLIQVADYPTIOLDSNLMQETI	413
354	QY	TGNPFLSLRPDENCOINRYGHPQMTITIVGELVENVLIQVADYPTIOLDSNLMQETI	413
416	QY	VTCKGATPEACTITISDPTCOIAQRVSPVAVDELCLSVRAFNFGSGTCYCNFTLGGD	475
416	QY	VTCKGATPEACTITISDPTCOIAQRVSPVAVDELCLSVRAFNFGSGTCYCNFTLGGD	475
416	QY	VTCKGATPEACTITISDPTCOIAQRVSPVAVDELCLSVRAFNFGSGTCYCNFTLGGD	475
414	QY	VTGQSGITPEVCTITISDPTCEIOTINVCYVDVDDMLCTVRRFTNGSGTCYCNFTLGGD	473
476	QY	ASLATSALSIIPGKDGLSPFLVNGVLSIGCLAMVMTVITLLYKHKHVKPIGNCTR	535
476	QY	ASLATSALSIIPGKDGLSPFLVNGVLSIGCLAMVMTVITLLYKHKHVKPIGNCTR	535
474	QY	STLTLSTLTLSPDOPASPLPMANSALSIWCCLAIEFTVITSLVLYKHKHVKPIGNCTR	533
536	QY	WYKSGKSLVFLSHQANRPFGRDQKDLQDK	566
536	QY	WYKSGKSLVFLSHQANRPFGRDQKDLQDK	566
534	QY	WYKSGKSLVFLSHQANRPFGRDQKDLQDK	566

```

RESULT 13
US-10-230-414-42
: Sequence 42, Application US/10230414
: Publication No. US20030050448A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Destrofers, Luc
: APPLICANT: Godeffroy, Michel
: APPLICANT: Godard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Smith, Victoria
: APPLICANT: Stephan, Jean-Philippe F.
: APPLICANT: Watanabe, Colin L.
: PRIORITY: 2002-04-09
: PRIORITY: 2002-04-09
: PRIORITY: 1997-09-09
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P35301C98
: CURRENT APPLICATION NUMBER: US/10/230,414
: CURRENT FILING DATE: 2002-08-28
: PRIOR APPLICATION NUMBER: 10/119,480
: PRIOR FILING DATE: 2002-04-09
: PRIOR APPLICATION NUMBER: 60/059113
: PRIOR FILING DATE: 1997-09-09
: PRIOR APPLICATION NUMBER: 60/062287
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/063549
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/064103
: PRIOR FILING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/069873
: PRIOR FILING DATE: 1997-12-12
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: NUMBER OF SEQ ID NOS: 246
: SEQ ID NO 42
: LENGTH: 572
- See File Wrapper or PALM.

```

```

:      TYPE: PRT
:      ORGANISM: Homo Sapien
US-10-230-414-42

Query Match          70.6%; score 2162; Db 9; Length 572;
Best Local Similarity 69.1%; Pred. No. 1.2e-170;
Matches 367; Conservative 78; Mismatches 87; Indels 12; Gaps
Qy      1 WSGVGVVLTLLAAGLPLOAKRPEFVLDLGHGEYDPDHNRNOLRGKSGSDENMDOLYP 60
Db      1 MECLVYFLGFLLAARLPDAAKRFHDVLGRPSAYNREHJNOLNGSSDNDNNKXLYP 60
Qy      61 VWRREGEKRWKSDGGRGVOAALTSDFALVGSNITFYVNIFFPRCKEDANGNIVYERNC 120
Db      61 WKRRGDWRKNSAGRGVOAVLTFSDPALVGSNITFAVNIFFPRCKEDANGNIVYERNC 120
Qy      121 RSDLSEADSPYYNNKTFGADDEDEWDTSGOHLRFPGKHPRDGHGPKKKNVYVFHTL 180
Db      121 RNEAGLSADPYNNKTAWSGDSGNGTGGSHNVFPGKHPRDGHGPKKKNVYVFHTL 180
Qy      181 GYOFKPGLOGCARSYNTNVNLTVPGVQMEVTVFRHRGATVPIISKVKVDVYITDQIPFV 240
Db      181 GYOFKPGLOGRVSRYNTNVNLTVPLQMLNEVTVFRHRGATVPIIAQVKVDVYITDQIPFV 240
Qy      241 TMFOYKGRSDSDRPFYRDLPIFDVLLHDHSPFLYKASISYKWNFGDNLGCVYSNNUTLN 300
Db      241 TMFOKDRNSDSETFKLDLPIFDVLLHDHSPFLYNTSYNYKWSFGDNLGLFVNTNHTVN 300
Qy      301 HTVYLVGTENFNLVQTAVTAVGCPSPPTSPSSSTSPSPASSSTPLST-----PSPSLMP 355
Db      301 HTVYLVGTFSLMLTAAAGAGCGCPPPPPPP-----PSKTFSLATLKSTSDNSNTPGP 353
Qy      356 GYKSKWELSDISNENCRINRYGVFPATITVDGLFVLYLVIOAVDPIPTLPQDNLMDPI 415
Db      354 TGQNPFLSRIPDENOCINRYGHFOATITVBSGLEVINITOMTDLVLMVPMFPSSSLDFV 413
Qy      416 VTKRGATPTEACTTIISDPTCOIAQRVYCSPVAVDELICLSVRFAFGSGTVGVNFILGDD 475
Db      416 VTQGSIGPEVTECTIISDPTCEITQNTVPSGVDFDDEKMLYVRFVNGSGTVCVNTLIGDD 473
Qy      476 ALSALTSALISIPGKDLGSLPRTVNGYSLTISGCLAMPVMTVILLYKKHKTKYPIGNCTR 535
Db      474 TSLAITSLSIYSDPDRPASPRLMANISALISVGCIAIFVTVISLLYVKKHKYNTIENSPG 533
Qy      536 NVYKRGKLSVLSHAKAPFSRGDEKDPLOOK 568
Db      534 NVYKSGLSVFLRANRATVFPNGDEKDPLOOK 566

RESULT 14
US-09-812-238B-2
; sequence 2; Application US/09812238B
; GENERAL INFORMATION:
; Patent No. US20020169132A1
; TITLE OF INVENTION: THERAPEUTIC ANTI-MELANOMA COMPOUNDS
; TITLE REFERENCE: GZ 2094.00
; CURRENT APPLICATION NUMBER: US/09/812.238B
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; SEQ ID NO 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-812-238B-2

Query Match          20.5%; score 627.5; Db 9; Length 661;
Best Local Similarity 27.1%; Pred. No. 1.3e-43;
Matches 176; Conservative 96; Mismatches 209; Indels 173; Gaps
Qy      7 VLVFLTLLAAGLPLOAKR--RFDVLDLGHGEYDPDHNRNOLRGWSSDNDMDOLYPPWRK 64

```

```

Db 9  LLHLAVIGALLAVKATVKPRNODMLG-----VSRQLR-----TKANNRLQIPENT- 53
Qy 65  GEGRKMSDEGGRVQALATSDSPALVGNITFVNLVFPKQCKEDANGNIVYERN-CRSD 123
Db 54  -EAPRLDWRGGOVSLKVSNDGPTLIGANASFIALNPFSGQVLVLPDGOVWNNITING 112
Qy 124  LELASDPYVYVWTKTGADDEDMDNTSOGQHLRPPDGKPPRPHGRKKNVYVHTLGOY 183
Db 113  SQVNGGQVYPOET--DDA-----CIEFDGGPCPSGSKASEVYVWTKWGY 159
Qy 184  FOKLGQCSARVSINTVNLTVGVOMEVIVFRHIG-RAYIPISKVKDYVITDQIPFVTM 242
Db 160  YOKNDRNSSDETFLRDLPIFFDVLHDPSPHFLNYSALSYKKNFGDNTGLFYSNNHFLNHT 302
Qy 243  YOKNDRNSSDETFLRDLPIFFDVLHDPSPHFLNYSALSYKKNFGDNTGLFYSNNHFLNHT 302
Db 220  SOLRALDGNKHFILNOLPLFALQIHDPSGYLAEDLSYTMDFGDSSTLISRALVYHT 279
Qy 303  YVLANGTFNFLTQVATP-----GPCP----- 324
Db 280  YLEPGVTAQVLAQAAIPLTSCGSPVPGTDDGHRPTAEAPNTAGOVPTTEWGTTPQ 339
Qy 325  SPTSPSSST-----SPSPASSPSPTLSTPSRLMP-----TG 357
Db 340  APTAESPCTTSVOVPTTEVISTAPVQNPMTAESTGMTPEKVPVSEVWGTTLAEMSTPEATG 399
Qy 358  YKSELSDI-----SNE----- 369
Db 400  MTPAEVSIVLGSCTTAQVTTWETTARELPPEPEGPDASSINSTESITGSLGPLLD 459
Qy 370  -----NCRINRYGFRATITIVDGLVNIQVADVPITLQPDNSLMDPI 415
Db 460  GTATRLVRKQVPLDCLVLRGSGFVTLIDVQGESAEILOA-----VPSGEGD-APELT 513
Qy 416  VTCKGATPEACTIISDPTCOIAQNRVCSVAVDELCLLSYVRAP-NGSGTYCVNFTLGD 474
Db 514  VSCOGGLPEACMEIISFGGCPQAPRLCQVLPSPACQVLVLAQILKGGSTTCLNLSLAD 573
Qy 475  DASLALTSALISIPCKDLG-SPLRTVNGVGLISGICLAMEVTWVITLYKK 523
Db 574  TNSLAVVSTOLIMPGQAGLGOVPLVIGIL-----LVLMVAVLASLIYRR 618

```

Search completed: April 7, 2003, 16:16:23
 Job time : 22 sec

RESULT 15

```

US-09-862-260A-2
; Sequence 2, Application US/09862260A
; Patent No. US2002008217A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: THERAPEUTIC ANTI-MELANOMA COMPOUNDS
; FILE REFERENCE: 126881210200
; CURRENT APPLICATION NUMBER: US/09/862.260A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/208.955
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/267.877
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-260A-2

```

Query Match 20.5%; Score 627.5; DB 10; Length 661;
 Best Local Similarity 27.1%; Pred. No. 1,3e-43;
 Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

```

Qy 7  VLIVFLLAGLQAAK-RFVDVIGHROYDPIHRENNOLRMSDSENEDSLYVWRR 64
Db 9  LLHLAVIGALLAVKATVKPRNODMLG-----VSRQLR-----TKANNRLQIPENT- 53

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 7, 2003, 16:08:12 ; Search time 21 Seconds
(without alignment)
2618.519 Million cell updates/sec

Title: US-09-943-075a-2

Perfect score: 3061
Sequence: 1 MESLGGVLLFLLAGLPLQ.....PFSKGRDKPLQDKPML 572

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2163	70.7	560	2 I38065	gene NMB protein - human
2	629	20.5	662	2 I38400	melanoma-associate
3	610	20.1	626	2 S53871	Pmel 17 protein -
4	610	20.1	626	2 S53871	melanocyte-specific
5	454	14.8	491	2 A4979	melanoma antigen h
6	133	4.3	868	2 A5420	cell differentiation
7	128.5	4.2	13055	2 T46590	hypothetical prote
8	126.5	4.1	611	2 S7621	hypothetical prote
9	125.5	4.1	351	2 S50754	hypothetical prote
10	123.5	4.0	279	2 T10361	hypothetical prote
11	123.5	4.0	473	2 S50755	hypothetical prote
12	121.5	4.0	422	2 T47901	endosperm specific
13	121.5	4.0	422	2 A3598	probable penicilli
14	120.5	3.9	620	2 A3598	internalin like pr
15	119.5	3.9	373	2 C70551	Clonase Ome
16	119.5	3.9	420	2 H84890	hypothetical prote
17	119	3.9	456	2 T17887	lysine/proline-ric
18	119	3.9	456	2 T17887	lin-1 protein - Ca
19	118.5	3.9	393	2 C33403	xylanase - Caldice
20	118	3.9	393	2 C33403	protein TPX-VT3
21	117.5	3.8	174	2 A62107	hydroxymethyl-ric
22	116.5	3.8	174	2 A62107	extensin homolog
23	114	3.7	474	2 S15921	proline-rich prote
24	114	3.7	1331	2 A48954	polycystic kidney
25	113.5	3.7	446	2 T07907	
26	111.5	3.6	221	2 T07176	
27	111.5	3.6	449	2 A24993	
28	111.5	3.6	544	2 T15447	
29	111	3.6	4302	2 A38971	

ALIGNMENTS

RESULT 1
138065
gene NMB protein - human
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
B:Accession: M38065 Subi, N.; van Dinter, I.M.; Degen, W.G.; van Muljen, G.N.; Rutter
Int. J. Cancer 60, 73-81, 1995
A:Title: nmb, a novel gene, is expressed in low-metastatic human melanoma cell lines
A:Reference number: I38065; MUID:95113576; PMID:7814155
A:Accession: I38065
A:Status: preliminary
A:Molecule type: mRNA
A:Accession: M38065
A:Cross-references: EMBL:X76534; NID:9656042; PID:CA54044.1; PID:9656043
C:Genetics: GDB:NMB
A:Gene: GDB:NMB
A:Cross-references: GDB:120237; OMIM:162340
A:Map position: 15q22-15qter

Query Match 70.7% Score 2163; DB 2: Length 560;
Best local similarity 69.4%; Pred. No. 1.7e+149;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;
Oy 1 MESLGGVLLFLLAGLPLQAAKRFVLDGQYDQJHWRNNOLRGSSDNDGDDLYP 60
Db 1 MECLYVFGFLGLLAARLPDAAKRFVDLGNERSPSATMREHNLQNGSSDNNKLLP 60
Oy 61 VMRRGGRKMDWEGGRVQAALSDSPALVGSNTFVNVLPFPCOCKDANGNIYVPRNC 120
Db 61 VMKGRDMKNSKNGKQVNLTSDSPPALVGSNTFAVNLIFPCOCKDANGNIYVPRNC 120
Oy 121 RSDLEASDPVYVNTTADDEEDNTSOGHLNFDCKKFFPRHGRKKNFYVPHL 180
Db 121 RNEAGLSADPVYVNTAWSEDSGNGTGGSHNVFPCKKFFPHHGRKKNFYVPHL 180
Oy 181 QYFCKLQACSAVSINTVNTLGPQVMEVIVFRHGRATPIISKVDVVTDTQIPFV 240
Db 181 QYFCKLQACSAVSINTVNTLGPQVMEVIVFRHGRATPIISKVDVVTDTQIPFV 240
Oy 241 TMYCKNDNSDDETLDDLPFFDFVLIHDSHFNLNYSALSYKWNFDCNLTGVSNNITLN 300
Db 241 TMYCKNDNSDDETLDDLPFFDFVLIHDSHFNLNYSALSYKWNFDCNLTGVSNNITLN 300
Oy 301 HTYVNGTFFNLTVOTAVPGRCPSPSPSSSTSPSPASSPSPSTLSPSPSLMPTGYS 360
Db 301 HTYVNGTFFNLTVOTAVPGRCPSPSPSSSTSPSPASSPSPSTLSPSPSLMPTGYS 360
Oy 361 MELSDLSNECNKRYFRATITVGGILEVNIIVQADVPIPTLPDNLMDFTVCKG 420
Db 361 MELSDLSNECNKRYFRATITVGGILEVNIIVQADVPIPTLPDNLMDFTVCKG 420
Oy 347 LELSRTPDENCINRYGHPQATITVGGILEVNIIVQADVPIPTLPDNLMDFTVCKG 406
Db 347 LELSRTPDENCINRYGHPQATITVGGILEVNIIVQADVPIPTLPDNLMDFTVCKG 406

serine rich pumil1
PEARL1 protein -
cellulase (EC 3.2.
protein PKA Ther
phosphatase
Clumping factor B
hypothetical prote
hypothetical prote
chitinase (EC 3.2.
N-acetylglucosamin
serine proteinase,
serine proteinase,
chitinase (EC 3.2.
lectin-like protai
mucin 2 precursor,

30 110.3 3.6 681 2 T39076
31 110.3 3.6 688 2 T07540
32 110.3 3.6 103 2 S72651
33 108 3.5 360 2 S72651
34 108 3.5 560 2 AD2389
35 108 3.5 877 2 F90070
36 107.5 3.5 602 2 AD2067
37 107 3.5 507 2 S05542
38 107 3.5 542 2 T31673
39 106.5 3.5 1036 2 T31673
40 106.5 3.5 629 2 D75390
41 106 3.5 629 2 D75390
42 105.5 3.4 224 2 T07861
43 105.5 3.4 287 2 S65765
44 105 3.4 582 2 T07952
45 105 3.4 3020 2 A43932

[illegible]

[illegible]

RESULT 9

S50754
 C:Accession: S50754
 C:Species: Chlamydomonas reinhardtii
 C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000

R:Woessner, J.P.; Molendijk, A.J.; van Egmond, P.; Kils, F.M.; Goodenough, U.W.; Haring, Plant Mol. Biol. 26, 947-960, 1994
 A:Title: Domain conservation in several volvoclean cell wall proteins.
 A:Reference number: S50754; MID:95093034; PMID:8000007
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-351 <OR>
 A:Cross-references: EMBL:L29028; NID:9530875; PID:AA853954.1; PID:g530878

Query Match 4.0%; Score 125.5; DB 2; Length 351;
 Best Local Similarity 42.3%; Pred. No. 0.14; 15; Indels 21; Gaps 4;
 Matches 33; Conservative 9; Mismatches 9;

Oy 296 NHTLN-----HTYVLNGTFN-----FNLT-VQTAVPGCPSPSPSPSSSS 335
 Db 128 NATVMTARAPTHLEFVYNNICDTRPAPYNCSTFNVTITPTSPSPSPSPSPSPS 187
 Oy 336 PSPASPSPT-LSTPSPS 352
 Db 188 PSPKASPSPKASPSPS 205

RESULT 10

T10361
 C:Accession: T10361
 C:Species: Orygia pseudotsugata nuclear polyhedrosis virus
 C:Date: 16-Jul-1999 #sequence_revision 18-Jul-1999 #text_change 21-Jul-2000

R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F. Virology 229, 381-399, 1997
 A:Title: The sequence of the Orygia pseudotsugata multinucleocapsid nuclear polyhedrosis virus
 A:Reference number: T10361; MID:97271300; PMID:9126251
 A:Status: preliminary
 A:Molecule type: DNA
 A:Superfamily: translated from GR/EMBL/DBS3

A:Residues: 1-279 <NR>
 A:Cross-references: EMBL:U75930; NID:9934903; PID:AA859091.1; PID:g1911338
 C:Superfamily: proline-rich protein 3

Query Match 4.0%; Score 123.5; DB 2; Length 279;
 Best Local Similarity 29.4%; Pred. No. 0.14; 34; Indels 39; Gaps 3;
 Matches 35; Conservative 11; Mismatches 39;

Oy 317 TAYVPGCPSP 375
 Db 40 TTPSP 92
 Oy 376 YGVFRATITVDGILEVNIQADVPIPTLPDNLMDFLVCKGATPCTACTILSDPT 434
 Db 93 -----PTPTSP 120

RESULT 11

S50755
 C:Accession: S50755
 C:Species: Chlamydomonas reinhardtii
 C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000

R:Woessner, J.P.; Molendijk, A.J.; van Egmond, P.; Kils, F.M.; Goodenough, U.W.; Haring, Plant Mol. Biol. 26, 947-960, 1994
 A:Title: Domain conservation in several volvoclean cell wall proteins.
 A:Reference number: S50754; MID:95093034; PMID:8000007
 A:Status: preliminary

A:Molecule type: mRNA
 A:Residues: 1-473 <OR>
 A:Cross-references: EMBL:L29029; NID:9530875; PID:AA853953.1; PID:g530876

Query Match 4.0%; Score 123; DB 2; Length 473;
 Best Local Similarity 26.2%; Pred. No. 0.21; 95; Indels 68; Gaps 16;
 Matches 68; Conservative 29; Mismatches 68;

Oy 126 LASQPVVYMMTGS-----ADDEDMDWTSSQOHLRFDDCKPEPRHGRKKNFVV 176
 Db 92 IEENRYISTYDTEKMGSPKVRDLKDW--VNAGSLVLDG--YSTASGN--TFVOL 144
 Oy 177 FH-TIAGVQFQLQO-CGARVSYNTVYLVTPGVMEVYVPRHGRAYTPSKVKVQVVVTD 234
 Db 145 IDLQV-----TKAGSGTGALYNGVN-----VYBR-ANSSSPGKTKTS----- 183
 Oy 235 QIFIVTYQKNDKNSDTEFLRDLFFOVLHDPHFNLAYSATSKNMFQDNTGLFVS 294
 Db 184 --PLVV--KGSRSRGITGCTSGAVLFSSNDKMTKAVTASAIT--MSVKGATIFG 236
 Oy 295 NNWTLNHTVYLVNG-----TFNNLVTVT-----NVGCPSTSPSPSPSPSP 338
 Db 237 SSFAMPH--LKGYEDMSGVAVTLNFRNTGASPSPKASPSKVPSPSPSPSPSPSP 293
 Oy 339 ASSESPSLTS--TPSPSLAP 355
 Db 294 KASPSPSPSPKASPSPSPS 313

RESULT 12

T47901
 C:Accession: T47901
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

R:Cholant, N.; Robert, C.; Brottler, P.; Mincker, P.; Cattolico, L.; Artiguenave, F.; W.; Rudd, S.; Lemcke, K.; Meyer, K.F.X. submitted to The Protein Sequence Database, March 2000
 A:Reference number: Z24479
 A:Accession: T47901
 A:Status: preliminary
 A:Molecule type: DNA
 A:Superfamily: 1-422 <CHO>
 A:Residues: 1-422 <CHO>
 A:Cross-references: EMBL:AL162295
 A:Experimental source: cultivar Columbia; BAC clone T4C21
 C:Genetics:

A:Map position: 3
 A:Note: T4C21.310

Query Match 4.0%; Score 121.5; DB 2; Length 422;
 Best Local Similarity 24.1%; Pred. No. 0.35; Indels 69; Gaps 8;
 Matches 47; Conservative 22; Mismatches 57;

Oy 251 SDPTF-LMDLPDIFDVLHDPHFNLAYSATSKNMFQDNTGLFVNNHPLNHTVYLVNGT 308
 Db 230 SDGARFAGVDTLNLTQREVSLLLEYHALAYK-----PKGSLTKNDALIS-LATNGA 283
 Oy 309 FNNPLTVQT-----AVPQPC 324
 Db 284 GRYELTSTSGDEVLTHTGVSRLADTVVDPTVYFTVONVLLPAELGCKSSSPAP 343
 Oy 325 SP--TPSPSSSTSP 370
 Db 344 EFVSAEPTTPKASPSPEAPSPAPSPAPSPAPSPAPSPAPSPAPSPAPSPAPSPAP 400

Oy 371 CRINRYGFRATITI 385
 Db 401 FHVNAFLPTALVTI 415

RESULT 13

T36588

probable penicillin-binding protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Accession: AF036598
 C:Query: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z1575

A:Accession: T36588

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-795

A:Cross-references: EMBL:AL049826; PIDN:CA842729.1; GSPDB:GN00070; SCOREDB:SCH24.23

A:Experimental source: strain A3(2)

C:Genetics:

C:Superfamily: penicillin-binding protein 1b

Query Match 4.0%; Score 121.5; DB 2; Length 795;

Best Local Similarity 21.9%; Pred. No. 0.83;

Matches 82; Conservative 53; Mismatches 134; Indels 105; Gaps 18;

QY 42 NOLGRSSDSEMEDEQLYVWVRGCRKMSGSRVGAALTSDFALVGSNITFVNVILV 101

DB 432 NOLKINNYDGSINENKGFQHQ-----ENDGQNDVGGPPYIMID- 472

QY 102 FPROCEDANGNIVYENCSDELASDPVYNTVGTGDEED-WEDNDSOGHLEPCK 160

DB 473 -REAMRESVSAFV-----QLGNDVGLQVKESAEDAGLLDSLAGAGYPSFSGV 522

QY 161 PFP---RPHG-----RKKRWYVVFHTLGO-YFQKLGQCSARVSINTVNLTVGP 205

DB 533 SPSALRMAGAYATPAGSKQKQEDYSVYKTKUGTUTLTKRVEKOATYKAVDVI--- 579

QY 205 QVKEVIVPRHG-RATVITSKV-----KDVYVI--TDQIPFTVWYO-KNDRNSS 251

DB 580 DVLTKVYKGTGSGNAQLGPRVAGKGTGTTGDNKSNAMFVYTPQLSTATSMYRPDDESIK 639

QY 252 DETFLRLPFIFFDLVHDPFLNYSALSYKKNFGDNTGLFVNNHTLNTVLTGTFNE 311

DB 640 DTFLEWCTGCKKILGAS-----FPS---TIWQDTHSDAKGM 676

QY 312 NL-----TWQTAVPQCPSPSPSS---TSPSPASSPSPULSTPSPSLMPTG 357

DB 677 KVEKFKPRIGEVINDAPSTPTTSPSPASSETEEVSPSPSPSDTQTSPP-PPAP 733

QY 358 YKSMELSDISNEMC 371

DB 734 TSCDFWDI---NC 744

RESULT 14

AG1598

Internalin like protein (IEXPG motif) [imported] - Listeria innocua (strain Clip11262)

C:Species: Listeria innocua

C:Accession: AG1598

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kraft, J.; Kunz, M.; Kunst, F.; Kuraphat, G.; Madueno, E.; Maltournam, A.; Ma

a; Title: Schuette; Simoes, A.; Torrez, A.;

A:Reference number: AB1077; NUID:21537279; PMID:11679669

A:Accession: AG1598

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-620 <GLA>

A:Cross-references: GB:AL592022; PIDN:CA96559.1; PID:gl6413801; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin1328

Query Match 3.9%; Score 120.5; DB 2; Length 620;

Best Local Similarity 27.4%; Pred. No. 0.7;

Matches 57; Conservative 23; Mismatches 73; Indels 55; Gaps 11;

QY 177 FHTLGGYFQKLGQCSARVSINTVNLTVGP-QVMKRVTFPRHGRAYTISKVKDVYVINDO 235

DB 391 ESTDGTGTVTLN-----ASNSVGLTADPQVQTVTV-----EQTVTIAE 429

QY 236 PFTVITVYOKNRNRSDETFRLDPIFFDLVHDPFLNYSALSYKKNFGDNTGLFVSN 295

DB 430 TEV---SYNNQKATAE-FLADINAVTN-----NSAKIT--TDTVTVDITAG 473

QY 296 NH-TLNTITVLTGNTFNLVQTVATVGPQPS-----PTPSPSSSTSPSPASSPPT 345

DB 474 EYTVTINAGTAKQKATPFKVTYKVVQDTPAPAPQPTPDPTSPDPPTTPPOASDPDPS 533

QY 346 -----LSTPSPSLMPTGYKSMELSDISN 366

DB 534 NQPAFAEDPGTSDSTG--SMDSTDPAN 559

RESULT 15

nodulin (clone GMenOD55-1) - soybean (fragment)

C:Species: Glycine max (soybean)

C:Accession: S37353

Ride Blank, C.; Mylona, P.; Yang, W.C.; Katinakis, P.; Bisseling, T.; Franssen, H.

plant Mol. Biol. 22, 1167-1171, 1993

A:Title: Characterization of the soybean early nodulin cDNA clone GMenOD55.

A:Reference number: S37353; NUID:94003074; PMID:8400132

A:Accession: S37353

A:Molecule type: mRNA

A:Residues: 1-137 <DEB>

A:Cross-references: EMBL:X69156; NID:g18584; PIDN:CAA48908.1; PID:g18585

C:Superfamily: plastocyanin

Query Match 3.9%; Score 119.5; DB 2; Length 137;

Best Local Similarity 56.8%; Pred. No. 0.44; 11; Indels 1; Gaps 1;

Matches 25; Conservative

QY 312 NLTVQTAVPQCPSPSPSSSTSPSPASSPPTLSTPSPSLMP 355

DB 68 NNTKKLLHSPSPS-SP 110

Search completed: April 7, 2003, 16:10:32

Job time: 30 secs

GenCere version 5.1.1.4_p5_4578
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 16:08:09 : Search time 14. Seconds

1694.604 Million cell updates/sec

Title: US-09-943-075a-2

Perfect score: 3061

Sequence: 1 MESGCVGVLLLAAGLPIQ.....PFSRGDEKPOLLDQKFWML 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.4

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2163	70.7	560	1 NMB_HUMAN	Q14956 Homo sapien
2	1463	64.5	762	1 P117_CHICK	G98317 Gallus gall
3	640.5	20.9	661	1 P117_HUMAN	P40967 Homo sapien
4	627.5	20.5	626	1 P117_MOUSE	Q60696 mus musculu
5	620	20.3	626	1 P117_BOVIN	Q06154 bos taurus
6	454	14.8	491	1 LAGC_DICDI	P42523 dictyostell
7	433	4.3	888	1 P117_MOUSE	P54383 dictyostell
8	431.5	4.0	592	1 G98317_CHICK	P54383 dictyostell
9	412.5	3.9	544	1 G98317_CHICK	Q06885 dictyostell
10	412.0	3.9	544	1 G98317_CHICK	Q06885 dictyostell
11	119.5	3.9	137	1 N551_SOYBN	Q05544 glycine max
12	118.5	3.9	555	1 GP1_CHLRE	Q9f9q6 glycine max
13	114	3.7	474	1 VTP3_TTVIV	P19275 thermoprote
14	114	3.7	1331	1 MANB_CALSA	P22533 caldocellum
15	113.5	3.6	409	1 G98317_CHICK	P22533 caldocellum
16	111.5	3.6	681	1 YDHE_SCHPO	Q92559 schizosacch
17	110.5	3.6	681	1 YDHE_SCHPO	Q92559 schizosacch
18	110	3.6	1036	1 GCT1_HUMAN	O15294 homo sapien
19	110	3.6	1039	1 GUNB_CALSA	P10474 c endogluc
20	109	3.5	360	1 VTP2_TTVIV	P19274 thermoprote
21	107.5	3.5	268	1 N020_MEDTR	P93329 medicago tr
22	107	3.5	316	1 P24_ENTFC	P56582 enterococcu
23	107	3.5	316	1 P24_ENTFC	P56582 enterococcu
24	105.5	3.4	324	1 Y091_NPVAC	P41479 autographa
25	105	3.4	5179	1 NUC2_HUMAN	Q02817 homo sapien
26	104.5	3.4	281	1 VAOA_VACCV	P29191 vaccinia vi
27	104.5	3.4	1742	1 GUNA_CALSA	P22534 caldocellum
28	103.5	3.4	626	1 GPBA_HUMAN	P07359 homo sapien
29	103.5	3.4	950	1 DC12_DROME	P18170 drosophila
30	103.5	3.4	1189	1 P18170_DROME	P18170 drosophila
31	103.5	3.4	1189	1 P18170_DROME	P18170 drosophila
32	103.5	3.4	2700	1 DC13_DROME	P18170 drosophila
33	103	3.4	738	1 VU47_HUSV6Z	P52549 human herpe

34	102	3.3	535	1 SPKC_SYNY3	P14745 synchocyst
35	101.5	3.3	1009	1 Y668_METTA	Q58645 methanococ
36	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
37	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
38	101.5	3.3	1569	1 YFJA_ECOLI	P52146 escherichia
39	101.5	3.3	1774	1 MSAS_PENPA	P22367 penicillium
40	100.5	3.3	792	1 PCNP_MOUSE	Q24h2 mus musculu
41	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
42	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
43	98.5	3.2	3317	1 CADN_RAT	P58365 ratus norv
44	98	3.2	682	1 AMPH_CHICK	P50478 gallus gall
45	98	3.2	734	1 G13B_DICDI	P34116 dictyostell

ALIGNMENTS

REMARK 1	
1 NMB_HUMAN	
2 NMB_HUMAN	
3 ID	560 AA.
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	

34	102	3.3	535	1 SPKC_SYNY3	P14745 synchocyst
35	101.5	3.3	1009	1 Y668_METTA	Q58645 methanococ
36	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
37	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
38	101.5	3.3	1569	1 YFJA_ECOLI	P52146 escherichia
39	101.5	3.3	1774	1 MSAS_PENPA	P22367 penicillium
40	100.5	3.3	792	1 PCNP_MOUSE	Q24h2 mus musculu
41	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
42	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
43	98.5	3.2	3317	1 CADN_RAT	P58365 ratus norv
44	98	3.2	682	1 AMPH_CHICK	P50478 gallus gall
45	98	3.2	734	1 G13B_DICDI	P34116 dictyostell

34	102	3.3	535	1 SPKC_SYNY3	P14745 synchocyst
35	101.5	3.3	1009	1 Y668_METTA	Q58645 methanococ
36	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
37	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
38	101.5	3.3	1569	1 YFJA_ECOLI	P52146 escherichia
39	101.5	3.3	1774	1 MSAS_PENPA	P22367 penicillium
40	100.5	3.3	792	1 PCNP_MOUSE	Q24h2 mus musculu
41	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
42	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
43	98.5	3.2	3317	1 CADN_RAT	P58365 ratus norv
44	98	3.2	682	1 AMPH_CHICK	P50478 gallus gall
45	98	3.2	734	1 G13B_DICDI	P34116 dictyostell

34	102	3.3	535	1 SPKC_SYNY3	P14745 synchocyst
35	101.5	3.3	1009	1 Y668_METTA	Q58645 methanococ
36	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
37	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
38	101.5	3.3	1569	1 YFJA_ECOLI	P52146 escherichia
39	101.5	3.3	1774	1 MSAS_PENPA	P22367 penicillium
40	100.5	3.3	792	1 PCNP_MOUSE	Q24h2 mus musculu
41	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
42	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
43	98.5	3.2	3317	1 CADN_RAT	P58365 ratus norv
44	98	3.2	682	1 AMPH_CHICK	P50478 gallus gall
45	98	3.2	734	1 G13B_DICDI	P34116 dictyostell

34	102	3.3	535	1 SPKC_SYNY3	P14745 synchocyst
35	101.5	3.3	1009	1 Y668_METTA	Q58645 methanococ
36	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
37	101.5	3.3	1333	1 UTX_MOUSE	Q58645 methanococ
38	101.5	3.3	1569	1 YFJA_ECOLI	P52146 escherichia
39	101.5	3.3	1774	1 MSAS_PENPA	P22367 penicillium
40	100.5	3.3	792	1 PCNP_MOUSE	Q24h2 mus musculu
41	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
42	99.5	3.3	1566	1 S422_HUMAN	P51531 homo musculu
43	98.5	3.2	3317	1 CADN_RAT	P58365 ratus norv
44	98	3.2	682	1 AMPH_CHICK	P50478 gallus gall
45	98	3.2	734	1 G13B_DICDI	P34116 dictyostell

```

DR PROSITE: P55093; PKD: 1.
KW Signal; Transmembrane; Glycoprotein; Polymorphism.
FT SIGNAL 22 560
FT DOMAIN 22 486 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 487 507 POTENTIAL.
FT DOMAIN 508 560 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 240 327 PKD.
FT DOMAIN 320 332 POLY-PKD.
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 195 S -> C (IN DBSNP:530436).
FT VARIANT 197 /FTID-VAR_012076.
FT SEQUENCE 560 AA; 62643 MW; 570035846CE3ECC CRC64;
Query Match 70.7%; Score 2163; DB 1; Length 560;
Best Local Similarity 69.4%; Pred. No. 2,6e-151;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

OY 1 MESLGCVLLVLLAGLPLOAKKFRFDVLGHQYDPDMRENNOLRWMSDENDEQLYP 60
DB 1 MCLLYEYLOFLVLLARLPIDAKNEFVLDLGNERSPATYMRHQNGNSSDENKELYP 60
OY 61 VWRGCRGRKMSGEGVQQAALTSQSPALVGSNITFVVLVFPQCKEDANGNTVYERNK 120
DB 1 RSDLELASDPVYNTGADDEKEDONTSCQHLRFQCKEDSPGHPGKYNVYVFTL 180
OY 121 RSDLELASDPVYNTGADDEKEDONTSCQHLRFQCKEDSPGHPGKYNVYVFTL 180
DB 121 RNEAGLSADPVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 180
OY 181 GYFKQKQCCSARVSNVNLVFLVQVMEVIVFRHGRAYVLSKVKDVTITDQIPFV 240
DB 181 GYFKQKQCCSARVSNVNLVFLVQVMEVIVFRHGRAYVLSKVKDVTITDQIPFV 240
OY 281 GYFKQKQCCSARVSNVNLVFLVQVMEVIVFRHGRAYVLSKVKDVTITDQIPFV 240
DB 281 TAYKQNDRSSDETFLEPLTFEVLVLDHPSFTLVAISYKKNKNGTCTGLVYSNHTLN 300
OY 241 TMEQNDRSSDETFLEPLTFEVLVLDHPSFTLVAISYKKNKNGTCTGLVYSNHTLN 300
DB 241 TMEQNDRSSDETFLEPLTFEVLVLDHPSFTLVAISYKKNKNGTCTGLVYSNHTLN 300
OY 301 HYVLNLTGFNLTQVATGRCPSSTFSSSTSPSPASSPSTPLSTPSPSLMPTQYS 360
DB 301 HYVLNLTGFNLTQVATGRCPSSTFSSSTSPSPASSPSTPLSTPSPSLMPTQYS 360
OY 301 HYVLNLTGFNLTQVATGRCPSSTFSSSTSPSPASSPSTPLSTPSPSLMPTQYS 360
DB 301 HYVLNLTGFNLTQVATGRCPSSTFSSSTSPSPASSPSTPLSTPSPSLMPTQYS 360
OY 361 MELSDINENKTRNGYVFNTHVLDGLVYLIVQAVDPIPLQDGLNSLMDFTVCKG 420
DB 361 MELSDINENKTRNGYVFNTHVLDGLVYLIVQAVDPIPLQDGLNSLMDFTVCKG 420
OY 347 LELSHIPDENCOINRYGHEQATITVIGELIVNLIQMTDVLMPVPWPESSLDIVETVCG 406
DB 347 LELSHIPDENCOINRYGHEQATITVIGELIVNLIQMTDVLMPVPWPESSLDIVETVCG 406
OY 421 ATPTAEATITSDPTQTAQNRVCSPVADEICLLSVRAFNGSTVYCNFTLGDGASLAL 480
DB 421 ATPTAEATITSDPTQTAQNRVCSPVADEICLLSVRAFNGSTVYCNFTLGDGASLAL 480
OY 407 SIPTVECTVITSDPTQTAQNRVCSPVADEICLLSVRAFNGSTVYCNFTLGDGASLAL 466
DB 407 SIPTVECTVITSDPTQTAQNRVCSPVADEICLLSVRAFNGSTVYCNFTLGDGASLAL 466
OY 481 TSAALSTPCGDLSPRTVNGVLSIGGLAMFVWTWILLKXKHKTYPKIGNTKRVYKG 540
DB 481 TSAALSTPCGDLSPRTVNGVLSIGGLAMFVWTWILLKXKHKTYPKIGNTKRVYKG 540
OY 467 TSTLCISVPDPPASPRLMANSALISVGLCATFTVVISLAVYKKIKETNPISNPGNVRS 526
DB 467 TSTLCISVPDPPASPRLMANSALISVGLCATFTVVISLAVYKKIKETNPISNPGNVRS 526
OY 541 KGLSVFSLSHAKAFPSRGDEKQPLQDX 568
DB 541 KGLSVFSLSHAKAFPSRGDEKQPLQDX 568
OY 527 KGLSVFSLSHAKAFPSRGDEKQPLQDX 554
DB 527 KGLSVFSLSHAKAFPSRGDEKQPLQDX 554

```

RESULT 2

Query Match

Best Local Similarity 49.8%; Pred. No. 2,7e-101;

Matches 280; Conservative 94; Mismatches 175; Indels 13; Gaps 3;

```

ONR_COTJA
ID QNR-71; 1998 (Rel. 36, Created)
DC 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE QNR-71 protein precursor.
GN QNR-71.
OS coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96283633; Pubmed=86700835;
RA Turquet N., Denhez P., Martin P., Planque N., Bailly M., Begue A.,
RA "Characterization of a new melanocyte-specific gene (QNR-71)
RA expressed in v-myc-transformed quail neuroretina.";
RL EMBL J. 15:3338-3350(1996).
RN [2]
RP REVISIONS.
RC TISSUE-Retina.
RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COULD BE INVOLVED IN MELANOGENESIS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: MELANOCYTE-SPECIFIC, RESTRICTED TO THE
CC -!- PIGMENTED LAYER OF THE RETINA AND THE EPIDERMIS.
CC -!- DEVELOPMENTAL STAGE: TRANSCRIPTIONALLY REGULATED BY MYC IN THE
CC -!- EXPRESSION PRECEDES MELANIZATION OF PIGMENTED EPITHELIAL CELLS.
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X94144; GenBank: U00409; RefSeq: NM_001004090; PKD_domain.
DR PROSITE: P55093; PKD: 1.
DR SMART: SM00089; PKD: 1.
DR PROSITE: P55093; PKD: 1.
KW Signal; Transmembrane; Glycoprotein.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 559 QNR-71 PROTEIN.
FT DOMAIN 483 507 POTENTIAL.
FT TRANSMEM 483 507 POTENTIAL.
FT DOMAIN 509 559 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 239 326 PKD.
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 559 AA; 61713 MW; 61E2793BD04N4YDD CRC64;

```



```
FT CONFLICT 274 274 L -> P (IN REF. 1 AND 5).
FT CONFLICT 587 587 P -> PVPGILLT (IN REF. 1).
FT CONFLICT 592 592 G -> GG (IN REF. 4).
FT CONFLICT 597 597 P -> R (IN REF. 1).
FT CONFLICT 642 661 SRKSCPCIGENSPLLSGQV -> ASSALVPLVRIAPSSVG
SO SEQUENCE 661 AA: 70255 MW: 84041FAB16715653 CRC64;

Query Match
Best Local Similarity 20.5%; Score 627.5; DB 1; Length 661;
Matches 176; Conservative 96; Mismatches 205; Indels 173; Gaps 19;

OY 7 LVVFLALLAGLPLOAK--RFRDVLGHEQYDPDMRENNLRGWSDEMDQGLYPWHR 64
DB 9 LHLAVLALLAGLAVKATKVPKRDMLG-----VSRQLR-----TRANROLYPMT- 53
OY 65 GEGRKDWSEGRVQVAAALTSDSPALGVSNITFVVLVFRPCKEDANGHYVERH-CRSD 123
DB 54 -EAQRLDCWRGGGVSLKVSNDPLTIGANASESIALNFGSGKVLPGQVIVWNTIING 112
OY 124 LEIASDPVYVWMTTGADEDEDNTSGQHILRFPPCKPFRPHGRKKNEVYVFTLQGY 183
DB 113 SQWGGQVYPOET---DDA-----CIPFGGRCPSGWSQKRSFYVWKTWGY 159
OY 184 FOKLGOCARSVSTNVLNLTVDGQVKNVYVFRIG-RATPISKVKOVVITQIPFVPM 242
DB 160 MQVLGPGVSGLSGTGRALGLTHFMVYVHREGSGSYPLAHSSSAITLDVFPFSVY 219
OY 243 YKNDNRSSDETLRDLPIFDVLIHDPSFLNYSATSKYKNQDNTGLFVSNITLHPT 302
DB 220 SOLRALDGNKHFRLKQPLTFAQLQHDPSGLAEADLSYTNDFGDSGSLISRALVYVHT 279
OY 303 YVLNCFNPNLTQVAP-----GCP----- 324
DB 280 XLEPGVTAQVLAQLTLTSCGSPVGTGDRTPAETAPNTAQVPTTEVGTTPGQ 339
OY 325 SPTSPSSST-----SPSPASSPILSTPSSILMP-----TG 357
DB 340 APTAPESGTVSVQVPTTEVISTAPVQMTAESTGMTPERKVPVSVGVMTTLAENSTPEATG 399
OY 358 YKSWELSDI-----SNE----- 369
DB 400 MTPAESVIVLISGTTAAQVTTETWETARELPEDPQDPDASIMSTSGSLGPLLD 459
OY 370 -----NCRINRYGFRATITVDGILEVNIIDVADVPITPLQDNLMDPI 415
DB 460 GTATLRVKRQVPLDCLVYRGSPVTLTDVIGIESAEILQA---VPSGEGD-APELT 513
OY 416 VTCKGATPTEACTITISPTQCIQANRCVSPVAVDELCLLSVRRAF-NGSGTCVYNLTLD 474
DB 514 VSCGGLPKAKHEISGFCQPAQRLQVLPSPQQLVHLQGLGSGTCLNLSLAD 573
OY 475 DASLALTSALISIPCKDLG-SPLTYNKNVLSICTLAHMYVTMLLYKK 523
DB 574 TNSLAVYSTQLIMQPEGAGLGQVPLVIGL-----LVLMVAVLASLIYVR 618

RESULT 5
PML7_MOUSE
ID PML7_MOUSE STANDARD; PRT; 626 AA.
AC Q60656;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Melanocyte protein Pmel 17 precursor (silver locus protein).
SR SILV OR PML17 OR D10H1253E OR S1.
OC Musculina (mouse).
OC Eukaria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Skin;
RC
```

```
FX MEDLINE=J5175358; PubMed=7870580;
RA Mon B.; Halaban R.; Ponnazhagan S.; Kim K.; Chintamaneni C.;
RT Mouse silver mutation is caused by a single base insertion in the
RT putative cytoplasmic domain of pmel 17.
RT Nucleic Acids Res. 23:154-158(1995).
RL -!- FUNCTION: COULD BE A MELANOGENIC ENZYME.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.
CC -!- DISEASE: DEFECTS IN SILV ARE THE CAUSE OF THE SILVER COAT COLOR
CC THE HAIR CYCLE BE DUE TO PREMATURE DEATH OF PIGMENT CELLS DURING
CC -!- SIMILARITY: BELONGS TO THE PML-17/NMB FAMILY.
CC -!- SIMILARITY: CONTAINS 1 PKD DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC to be construed as an endorsement of the product or service of any
CC entities mentioned in this statement is not removed. Usage by for commercial
CC or send an email to license@ebi.ac.uk. (See http://www.isb-sib.ch/announce/
CC
CC EMBL: U14133; AAA69538.1;
DB MGD: MGI:98301; si.
DR InterPro; IPR006001; PKD_domain.
DR Pfam; PF08081; PKD; 1.
DR SMART; SM0089; PKD; 1.
DR PROSITE; PSS009; PKD; 1.
DR TrEMBL; P05095; PKD; 1.
KW Transmembrane glycoprotein; Signal; Melanin biosynthesis; Repeat;
KW Disease mutation.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 626 MELANOCYTE PROTEIN Pmel 17.
FT DOMAIN 25 562 EXTRACELLULAR (POTENTIAL).
FT TRANSHEM 563 583 POTENTIAL.
FT DOMAIN 584 626 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 584 592 PKD.
FT DOMAIN 592 626 1 X 13 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 315 327 1.
FT REPEAT 328 340 2.
FT REPEAT 341 353 3.
FT REPEAT 354 366 4.
FT REPEAT 367 379 5.
FT REPEAT 380 392 6.
FT REPEAT 393 411 7.
FT CARBOHYD 406 407 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 535 535 N-LINKED (GLCNAC. ) (POTENTIAL).
FT VARIANT 170 170 S -> L (IN SILVER).
FT VARIANT 175 175 R -> G (IN SILVER).
FT VARIANT 373 373 F -> N (IN SILVER).
FT VARIANT 471 471 F -> S (IN SILVER).
FT VARIANT 603 626 AAPASGLRARGLGENSPLLSGQV -> SSASURSSRPW
FT REPEAT 603 626 AAPASGLRARGLGENSPLLSGQV (IN SILVER).
SO SEQUENCE 626 AA: 65980 MW: 748941D2B3F1044 CRC64;

Query Match
Best Local Similarity 20.3%; Score 620; DB 1; Length 626;
Matches 174; Conservative 91; Mismatches 198; Indels 164; Gaps 19;

OY 54 WDEQLFPVWRGGERKDWSEGRVQVAAALTSDSPALGVSNITFVVLVFRPCKEDANGN 113
DB 44 WNRGQVPEKTEVGS--SNCRGGGVSLVRIINDOPLVGNAMSPSIALHIFGSKQVLPDQ 101
OY 114 IVYERINCRSDLELASDPVYVWNTTGADEDEDNT---SQ---GQHLR-----pp 157
DB 102 VT-----WANNITINGSQVNGGQVPIVPOEPDDACVFP 133
OY 158 DCKPFRPHGSKWHPVVFHTLQGYFQKLGOCARSVSTNVLNLTGVQVMEVYVFRIG 217
DB 134 DQGLPSPGPKPRKRSFYVWKTWKTWQVLGGPVYRSRSSTRAHKLGHMTVETVYHRRG 193
```

```

QY 218 -RAYIPISKVDVITVDQIPFVPMYKOKNDKNSDDETFDLRDLPIFDVLHDPHSFLNY 276
Db 194 SQSVPLAHASSTFTTIDOVDFSVGSQQLACDETKHFLRNHPLIFALQLHDPGSLAE 253
QY 277 SALSIXKNFNDGTLGVSNHNTLNHTYTLNLTGTFNMLVQVAVP-----GPCSP- 326
Db 254 ADLSTVWDFGDTGLTSLRDLVHTYLGSVTSVAQVQLNAIPLVSCGSSVPVPTTDCY 313
QY 327 -----TPSPSSPS-----PASSPSPTLS----- 347
Db 314 MPTAEAPOTTSRQGTITKVGTPFQOMETTPQSGTTVVMQPTTEVATISEQMLTSAVID 373
QY 348 -----TPS-PS-----LMPTGYKSMELSDISNE----- 369
Db 374 TILAESTVTEGTGTTTPRPSGTTVAQATTEPGDAPGLLPQSGTSLSPLLDDTDTML 433
QY 370 -----NCIRNRYGPRATTIVDGLILEVNILOVADNPITPLQDNLMDPIVCKGAT 422
Db 434 VKROVPLDVLVRYGFSPLADIVOGIESAELQA-VFP-----SEGDAPFELVSCGGL 487
QY 423 PTEACTISDPTCGIAONRVCSPVAVDELCLLSVRRAF-NGSGTYCVNFTLGDASLALIT 481
Db 488 PKACMSDSCGPTQACRLQCSVPSPDCCQLVHLQVKGSGTCLNLSADNLSLAVA 547
QY 482 SALTIPCKDLG--SEPLTYGVNGSLVIGCLAMEVNTLLYKHYKPIGNTGRNVV 538
Db 548 STOLVVPQDGLQAPL-----LVGLLLVAVVLASLLGLDLSRAQEPKCHM--- 598
QY 539 KKGKLSVFLSHAKAPFSRDXREKDPIL 565
Db 599 -----VALTNAPASGLRAGLGENSPILL 621

RESULT 6
ID PM17_BOVIN STANDARD; PRT; 491 AA.
AC Q06154;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanocyte protein Pmel 17 (Retinal pigment epithelial-specific
GN Melanocyte protein Pmel 17 (Retinal pigment epithelial-specific
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC ISSUES:
RC MIM:6089321263; PubMed=1478275;
RA Kim R.Y., Wistow G.J.;
RT The cDNA Rpe1 and monoclonal antibody HMB-50 define gene products
RT preferentially expressed in retinal pigment epithelium.*;
RL Exp. Eye Res. 55:657-662(1992).
CC -/- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -/- TISSUE SPECIFICITY: RETINAL PIGMENT EPITHELIUM.
CC -/- SIMILARITY: PHE17/PHE17/MBE FAMILY.
CC -/- SIMILARITY: CONTAINS 1 PKD DOMAIN
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to: license@isb-sib.ch)
-----
DR EMBL: M81103; AAA30419.1;
DR InterPro: IPR000601; PKD_domain.
DR Pfam: PF00801; PKD; 1.

```

```

DR SMART: SM00089; PKD; 1.
DR PROSITE: PS50093; PKD; 1.
KW Transmembrane; Glycoprotein; Melanin biosynthesis; Repeat.
FT NON_TER 1
FT DOMAIN 41 423
FT DOMAIN 42 443
FT DOMAIN 43 463
FT DOMAIN 44 491
FT DOMAIN 45 500
FT DOMAIN 60 150
FT DOMAIN 148 256
FT REPEAT 148 160 1.
FT REPEAT 161 173 2.
FT REPEAT 174 186 3.
FT REPEAT 187 199 5.
FT REPEAT 200 212 6.
FT REPEAT 213 225 7.
FT REPEAT 226 243 8.
FT REPEAT 244 256 8.
FT DOMAIN 304 394
FT CARBOHYD 269 269
FT CARBOHYD 396 396
FT SEQUENCE 491 AA; 51669 MW; 2BFE5DFD397D6D CRC64;
SO QUERY MATCH
Query Match 14.88; Score 454; DB 1; Length 491;
Best Local Similarity 26.18; Pct. Mismatches 137; Indels 132; Gaps 11;
Matches 120; Conservative 71; NMatches 137;
QY 182 QYFORKLGGCCARVSINVTNLVGPQWMEVIVFRRHG-RAYIPISKVDVITVDQIPFV 240
Db 1 QYVQVGLGPGVGLSGICTOKAMLCITNMETVYTHRGSSVYPLANSSAFFITIDVPSV 60
QY 241 TAYOKNDKNSDDETFDLRDLPIFDVLHDPHSFLNYSALSIVKMNKGDTGLFVSNHNLN 300
Db 61 SVSLOALDQNRKRLKOPLTALQLHDPSCYLACADLSYTWDRGSTGLTSLKALITV 120
QY 301 HTYVLNLTNFNLVQTAVP-----GPCSP- 326
Db 121 HTYLESGPTVAQVQLNAIPLTSCGSSVPVGTDRHWTAEPACTAGVPTTEWGTPP 180
QY 327 -----TPSPSSPSPTLSTESP-----SLMPTGYKSMELSDISNE----- 369
Db 181 GQVPTAEAPCTTGWVPTTEVGTTPPEQVATSKVLSITTVEMPTAKATGRTEVSTTSPS 240
QY 332 SS-----TPSPSSPSPTLSTESP-----SLMPTGYKSMELSDISNE----- 369
Db 241 GTTVTQGTTELPVETAGVETSEFAGTSNFSFMTPTAGTSLSPDDTATLVLEKRA 300
QY 370 --NCIRNRYGPRATTIVDGLILEVNILOVADNPITPLQDNLMDPIVCKGATTEAC 427
Db 301 PLDCLVLYRYGFSPLDIVS-----IESAEILLQAVSSSGDAFELTVSCQGLPKAC 353
QY 428 TITSDPTCGIAONRVCSPVAVDELCLLSVRRAF-NGSGTYCVNFTLGDASLALIT 486
Db 354 MDLSSPGCQIAPRLQCPVPSPACQLVHLQVKGSGTCLNLSADNLSLAVMSTOLV 413
QY 487 IPCKDLG---SPLTYGVNGSLVIGCLAMEVNTLLYK 523
Db 414 MPQDENGRLQAPLV-GILLVTLA-----LLASLIYER 446

RESULT 7
ID LAGC_DICDI STANDARD; PRT; 888 AA.
AC P42523;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE LAGC aggregate C protein precursor.
GN LAGC
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP SEQUENCE FROM N.A.

```

Query Match 4.3%; Score 133; DB 1; Length 888;
Best Local Similarity 20.68; Pred. No. 0.046;
Matches 91; Conservative 59; Mismatches 188; Indels 104; Gaps 18;

147	--	I-GCGPNFLINSIDVKKFPYKGNFSDS----	PDCNITVTTP-----PSSGPKFLPYDGTG	199
320	PGICPEPTSPSPSSSTPSPPASPSPLTSTS----	PSLMPFYGSME----	LSDSNENC	371
360	DNFVPFESTPSISSVSDDSKQIITINGDNFTDKD-VKVFSGDIGNPFISVNHQKI	259		
372	RINKY-----GYPRATIVTGILAVNI IOAVDPV IPTPLDSNSLMDFIVTCAGATPFA	426		
200	OYNHNVSDCTGPMYSNTIVDESVIEKNY IHCFPAVITISLVNSNNVDFIVTCAGATPFA	426		
260	QYHNVDQPCQNRKSVTVDESVIEKNY IHCFPAVITISLVNSNNVDFIVTCAGATPFA	426		
427	CITLIISNDOTCAQRNGSPVADEL-GLLSVRFNFGSGTVCVNFLTIGDGDAISALTSALI	385		
320	LNLYLTPTSITGDVKYIKASTWTECKLEKDANEJGGKVLNPNVFFGCCDSTPNGVSFTY	379		
486	SIPKGSLGSLPRLVNGVNLISIG	507		
380	NIPTLSRGS----YSGNLVTLIG	398		
<hr/>				
RESULT 8				
DB	AC	GUN1-ACE	STANDARD:	PRT: 562 AA.
DB	AC	Q54583	1995 (Rel. 34, Created)	
DB	AC	01-OCT-1995 (Rel. 34, Last sequence update)		
DB	DE	01-OCT-1995 (Rel. 34, Last annotation update)		
DB	DE	Endoglucanase E1 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase E1)		
DB	DE	(Cellulase E1) (Endocellulase E1).		
DB	DE	Acidothermus cellulolyticus.		
DB	CC	Bacteria; Actinobacteridae; Actinobacteriales (class); Actinobacteridae;		
DB	CC	Actinomycetaceae; Streptomyces; Frankineae; Acidothermaceae; Acidothermus.		
DB	CC	[1] NCBI Tax ID=28049.		
DB	RP	SEQUENCE FROM N.A.		
DB	RP	STRAIN=ATCC 43068 / 11B;		
DB	RC	Laymon R.A., Himmel M.E., Thomas S.R.;		
DB	RC	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.		
DB	RR	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 42-398.		
DB	RR	MEDLINE=16340558; PubMed=8718654;		
DB	REX	Sakon J., Adney W.S., Himmel M.E., Thomas S.R., Kaplus P.A.;		
DB	RT	"Crystal structure of thermostable family 5 endocellulase E1 from		
DB	RT	'Acidothermus cellulolyticus in complex with cellobiose';		
DB	RT	Biochemistry 35:10648-10660(1996).		
DB	CCC	-FUNCTION: THERMOSTABLE ENZYME WITH AN OPTIMAL TEMPERATURE OF 81		
DB	CCC	DEGREES CELSIUS.		
DB	CCC	ENZYME CLASSIFICATION: EC 3.2.1.4, CELLULOSE		
DB	CCC	HYDROLASES.		
DB	CCC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic		
DB	CCC	linkages in cellulose.		
DB	CCC	-1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL		
DB	CCC	HYDROLASES).		
<hr/>				
THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL Data Bank				
and the Wellcome Trust. The database is available free of charge to all				
users by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See http://www.isb-sib.ch/announce/				
or send an email to licenselab@sib-sib.ch).				
<hr/>				
DR	BL	U02214	AAN5477.1;	
DR	DR	InterPro: IP00015476		
DR	DR	Inter-Pro: IP0001919; Bac_cellose-blind		
DR	DR	Inter-Pro: IP0001547; GH_5		
DR	DR	Pfam: PF00150; cellulase; 1.		
DR	DR	Pfam: PF00553; CBM_2; 1.		
DR	DR	PROSITE: PS00659; GLYCOSYL-HYDROL_F5; 1.		
KW	CC	Cellulase degradation; Hydrolyase; Glucosidase; Signal; 3D-structure.		
FT	FT	SIGNAL	1 41	ENDOGALACTAMINASE EL.
FT	FT	DONAT	42 500	CATALYTIC
FT	FT	DONAT	401 461	PRO/SER/THR-RICH (LINKER).
FT	FT	DONAT	462 562	CELLULOSE-BINDING (BY SIMILARITY).

```

FT ACT_SIZE 203 203 PHOTON DONOR.
FT ACT_SIZE 163 163 NUCLEOPHILE.
FT DISULFID 75 163
FT DISULFID 209 212
SQ SEQUENCE 562 AA: 60747 MW: 84E6256406A35041 CRC64:

Query Match 4.3%; Score 131.5; DB 1; Length 562;
Best Local Similarity 20.9%; Pred. No. 0.033;
Matches 63; Conservative 51; Mismatches 84; Indels 103; Gaps 13;

OY 230 YVTDQI-PPIYTWKNDKNSDFTFLRDLFPDFVLHPSHFLNYSALSKWNP-- 285
DB 310 YLFNPIAPVWLGEFGTTLQSTDTQTKLTKL-----VOYLRLPT--AQYGADFQKTEWSM 362

OY 286 ----GNTGLFVSNHNLINHTVYLVNGTFNLTUOTPAV-----PGPCSPFP 328
DB 363 NPDGDTGGILKUDQVWD--TVADG---TLAPKISIFDFVGASPSQSFVSFVSP 417

OY 329 SPSFSTSPSPASSPSTLSPSPGMPYGYKMSLDSINENCRINRYGYFRATITVDG 388
DB 418 SPSFASRTPPRPFPPTAS-PPITPLTPTA----- 445

OY 389 ILEVNIIQVADPIPTLPDNLDMDFIVTCKGATPEACTIISDPTQCIQAHRVCSPIV 448
DB 446 -----TPPI-----ASTPSTPAAGCACTASYO-----VNS 473

OY 449 DELCLLSVRAPNGSG---TYCVNFTLGDGASIALT-SALISTPKGLDKSLPLRTVNGV 502
DB 474 DWNGFTVTVAVTNSGVSATKTWTYSWTFGNGQITITNSMNAAVTQNGQSVTARNNSYNNV 533

OY 503 I 503
DB 534 I 534

RESULT 9
ID Y091_NPVP STANDARD; PRT; 279 AA.
AC O10341.1; 1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 29.3 kDa protein (ORF92).
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpNPV).
SC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
CC Nucleopolyhedrovirus.
EMBL: U75930; AAC59091.1; -.
NCBI_TaxID:164623;
RN 1[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91271300; PubMed=9126251;
RA Ahrens C.H., Russell R., Funk C.J., Evans J., Harwood S.,
RA Rohmann G.F.;
RA "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome."
RT PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES 90:1166-1170
CC 1- SIMILARITY: TO CORRESPONDING ORF IN ACHNPV.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04286; AAC37369.1; -.
DR D12cyob; D002035; gppa.
FM Orygia pseudotsugata.
FT STORL 20 544
FT CHAIN 20 544
FT DOMAIN 20 489
FT TRANSMEM 490 510
FT DOMAIN 511 544
FT DOMAIN 117 208
FT CARBOHYD 80 80
FT CARBOHYD 208 244
FT CARBOHYD 308 308
FT CARBOHYD 332 332
FT CARBOHYD 366 366
FT CARBOHYD 380 380
FT CARBOHYD 410 410
FT CARBOHYD 422 422
FT CARBOHYD 478 478
SQ SEQUENCE 544 AA: 59156 MW: 95430DB7C7D0A0 CRC64:

Query Match 3.9%; Score 120; DB 1; Length 544;
Best Local Similarity 19.2%; Pred. No. 0.22;
Matches 85; Conservative 61; Mismatches 105; Indels 192; Gaps 25;

OY 224 SKVQVYVTDQIPIYTWKNDKNSDFTFLRDLFPDFVLHPSHFLNYSALSKW 283
DB 21 SNKKD-----DQSPETVKI-----NSNDH-----LPLVYEIII--SGHF-----SI----- 54

OY 284 NEGDTNGTLFVS-----NNHTLNT-----YLVNGTF 309

```

```

OY 317 TAVPGCPSTPSPSSPSPASSPSTLS-TPSPSLMPTGYKMSLDSINENCRINR 375
DB 40 TTPSPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 92

OY 376 YGVFRATITVDGILEVNIQVADPIPTLPDNLDMDFIVTCKGATPEACTIISDPT 434
DB 93 -----PPTTSPSPS-----PPTTSPSPSPTPTPTPTPTPTPTPTPTPT 120

RESULT 10
ID GP10_DICDI STANDARD; PRT; 544 AA.
AC Q06885;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Glycoprotein GP100 precursor (P29F8)..
OS Dictyostelium discoideum (Slime mold).
CC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID:44689;
RN 1[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94124612; PubMed=8294503;
RA "Bach A., Mueller, T., Taudenberger A., Taranto P., Gerisch G.;
RT glycoprotein of D. discoideum by a transmembrane region does not
RT impede cell adhesion but reduces residence time on the cell
RT surface."
RL J. Cell Biol. 124:205-215 (1994).
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- DEVELOPMENTAL STAGE: RESTRICTED TO THE AGGREGATION STAGE OF
CC ADHESION.
CC 1- PTM: N- AND O-GLYCOSYLATED.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04286; AAC37369.1; -.
DR D12cyob; D002035; gppa.
FM Orygia pseudotsugata.
FT STORL 20 544
FT CHAIN 20 544
FT DOMAIN 20 489
FT TRANSMEM 490 510
FT DOMAIN 511 544
FT DOMAIN 117 208
FT CARBOHYD 80 80
FT CARBOHYD 208 244
FT CARBOHYD 308 308
FT CARBOHYD 332 332
FT CARBOHYD 366 366
FT CARBOHYD 380 380
FT CARBOHYD 410 410
FT CARBOHYD 422 422
FT CARBOHYD 478 478
SQ SEQUENCE 544 AA: 59156 MW: 95430DB7C7D0A0 CRC64:

Query Match 3.9%; Score 120; DB 1; Length 544;
Best Local Similarity 19.2%; Pred. No. 0.22;
Matches 85; Conservative 61; Mismatches 105; Indels 192; Gaps 25;

OY 224 SKVQVYVTDQIPIYTWKNDKNSDFTFLRDLFPDFVLHPSHFLNYSALSKW 283
DB 21 SNKKD-----DQSPETVKI-----NSNDH-----LPLVYEIII--SGHF-----SI----- 54

OY 284 NEGDTNGTLFVS-----NNHTLNT-----YLVNGTF 309

```

Db 55 ---DKTGLVTVYLGDKIHDFEIKGVGLNVTGEPONNPIPTVTSINPDQNDREVSQSS 111
 Qy 110 NFNLVOTAVGPCSPSPSSSTS : : : : :
 Db 312 DESKQTPPTPTTKSKPTSTPTSPQITPTPTVSPVPTQPTSKSTPTSTPTPTTP 171
 Qy 345 TLTSPSPS- : : : : :
 Db 172 T-STPTPTSPPTPTPTTKSKMPTKPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 218
 Qy 383 ITVDGILEWNIOWADVPIPL : : : : :
 Db 219 -----DLIFNFIOTDIELENGKITSYNGQSSKHQFNPSIKSTERTSYGVGVFKG 273
 Qy 421 A-----TPEAC-TIETDPT---COIAQNRVCS-PAVDEICLLSVRAFPNS 463
 Db 274 SFYTDDEYSKIGKNGKICETTSSTRCYITNGTGGCYITIDNL-LNFIDNNGNS 332
 Qy 464 GYVC-----VNFGLD-DASIALTSALTSIKSGVGLSFLRVNGVLSI : : : : :
 Db 333 LTYCVNPIIDKIVGDKKL-----TLITIGNFS-----LNNATVIEKPNKNCRS 382
 Qy 510 AMEVMTWTLVLYKKHKTYPION 532
 Db 383 NVLLSTDTLIFICELSKSTHTPLSS 405
 RESULT 11
 N551_SOYBN STANDARD; PRT: 137 AA.
 AC 005544;
 DT 15-DEC-1998 (Rel. 37, Created)
 DE 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-1999 (Rel. 38, Last annotation update)
 DE Early nodulin 55-1 precursor (N55-1) (Fragment).
 OS ENO055-1.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 RN [1]_TaxID=3847;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Williams; TISSUE=ROOT;
 RA MEDLINE=94003074; PubMed=8400132;
 RX de Blenk C., Wylona P., Rafanek P.C., Bisseling T., Franssen H.;
 RT Characterization of the soybean early nodulin cDNA GENOD55-1;
 CC 1- SUBCELLULAR LOCATION: PERIPLASMIC MEMBRANE (POTENTIAL).
 CC DEVELOPMENTAL STAGE: EXPRESSED AT EARLY STAGES OF NODULE
 CC PLANTS AFTER WHICH LEVELS DECREASE.
 CC -1- INDUCTION: DURING NODULATION IN LEGUME ROOTS AFTER RHIZOBIUM
 CC INFECTION, AND AFTER RELEASE OF BACTERIA FROM THE INFECTION
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC emil@isb-sib.ch).

FT CHAIN ? 137 EARLY NODULIN 55-1.
 FT DOMAIN ? ? PLASTOCYANIN-LIKE.
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 137 AA; 14857 MW; 988DE6F95A2EB0 CRC64;
 Query Match 3.9%; Score 119.5; DB 1; Length 137;
 Best Local Similarity 56.8%; Pred. No. 0.038;
 Matches 23; Conservative 7; Mismatches 11; Indels 1; Gaps 1;
 Qy 312 NUTVOTAVGPCSPSPSSSTSFPASSPPTLSTPSPSLAP 355
 Db 68 NWTAKKILNSPSPS-SFSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 110
 RESULT 12
 ID GPI_CHLRE STANDARD; PRT: 555 AA.
 AC 059367;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vegetative cell wall protein gpl precursor (Hydroxyproline-rich
 DE glycoprotein 1).
 GN GPI.
 OC Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadales; Chlamydomonadales.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21159092; PubMed=11258910;
 RX Ferris P.J., Woessner J.P., Waffenschmidt S., Kilz S., Drees J.,
 RT Glycosylated polyproline II rods-with-kinks as a structural motif in
 RT plant hydroxyproline-rich glycoproteins.*,
 RL Biochemistry 40:2978-2987(2001).
 RN [2]
 RP PARTIAL PRELIMINARY SEQUENCE FROM N.A.
 RA MEDLINE=51017504; PubMed=1695225;
 RX "Cell Wall" Section in Chlamydomonas: accumulation of m6nas
 RT encoding cell wall hydroxyproline-rich glycoproteins.*,
 RL proc. Natl. Acad. Sci. U.S.A. 87:7355-7359(1990).
 CC -1- FUNCTION: Major component of the outer cell wall w6 (crystalline)
 CC layer.
 CC -1- SUBUNIT: Associates with GP2 and GP3.
 CC -1- PTM: Glycosylated and O-glycosylated.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC emil@isb-sib.ch).

EMBL: AF309494; AAG45420.1;
 EMBL: M58496; AAG69706.1; ALT_SEQ.
 DR GlycoSuiteDB; O9FP06;
 DR InterPro; IPR002965; P-rich_extensin.
 DR InterPro; IPR003882; Pistill_extensin.
 DR PRINTS; PR01217; PRICHEXTENS.
 DR PROSITE; PS01218; PESTLEXTENSIN.
 KW GLYCOPOLYSACCHARIN, Repeat, Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 555 VEGETATIVE CELL WALL PROTEIN GPI.
 FT DOMAIN 40 339 49 X 5 AA APPROXIMATE PPSXP REPEATS.
 FT DOMAIN 259 279 POLY-PRO.
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SIGNAL <1 ? POTENTIAL.

DT	01-AUG-1991 (Rel. 19, Created)
DE	11-JUL-1993 (Rel. 26, Last update)
DE	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-
DE	beta-mannosidase A (EC 3.2.1.78) (Beta-mannanase) [Endo-1,4-
DE	mannanase] Endo-1,4-beta-glucanase (EC 3.2.1.4) (Cellulase)].
MA	MANA.
GN	Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
OS	Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
OC	Caldocellales; Caldocellaceae; Caldocellaceae; Caldocellaceae;
CC	NCBI_TaxID=44001.
KP	[1]
RN	SEQUENCE FROM N.A.
RX	MEDLINE=93119139; PubMed=1476429;
RX	Gibbs M.D., Saul D.J., Luthi E., Bergquist P.L.;
RT	'The beta-mannanase from "Caldocellum saccharolyticum" is part of a
RT	multidomain enzyme.';
RT	[2]. Environ. Microbiol. 58:3864-3867(1992).
RN	SEQUENCE OF 1-346 FROM N.A.
RX	MEDLINE=G1247819; PubMed=2039230;
RX	Luehl E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
RT	"Cloning, sequence analysis, and expression in Escherichia coli of a
RT	gene coding for a beta-mannanase from the extremely thermophilic
RT	bacterium "Caldocellum saccharolyticum".
RT	[20]. J. Biol. Chem. 267:9090-9093(1992).
CC	-1- FUNCTION: DEGRADATION OF HEMICELLULOSES, THE SECOND MOST ABUNDANT
CC	POLYSACCHARIDES IN NATURE. CONTAINS TWO CATALYTIC DOMAINS WITH
CC	MANNAASE AND ENDOGLUCANASE ACTIVITIES.
CC	-1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic
CC	linkages in mannans, galactomannans, glucomannans, and
CC	galactoglucomannans. Endohydrolysis of 1,4-beta-D-glucosidic
CC	linkages in cellulose.
CC	-1- MISCELLANEOUS: THIS ENZYME IS MOST ACTIVE AT PH 5 AND 80 DEGREES
CC	CELSIUS.
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
CC	A (FAMILY 5 OF GLYCOSYL HYDROLASES).
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
CC	A (FAMILY 44 OF GLYCOSYL HYDROLASES).
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by for commercial
CC	entities requires a license agreement. (See http://www.isb-sib.ch/commercial/
CC	it send an email to license@isb-sib.ch).
DR	EMBL: L01257; AAAT1887.1; *
DR	EMBL: M30631; AAA72861.1; *
DR	PIR: B43745; B43745.
DR	PIR: A48954; A48954.
DR	HSPR: Q06851; INBC.
DR	InterPro: IPR001549; CB_D_3.
DR	Gene: G12478; G12478.
DR	Pfam: PF00150; cellulase5.
DR	Pfam: PF00942; CBM_3; 1.
DR	ProDom: PD001947; CB_D_3; 2.
DR	ProSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KK	Hydrolases; Glycosidase; Cellulose degradation; Signal;
KW	Multifunctional enzyme.
FT	CHAIN 42 1331
FT	DOMAIN 42 325
FT	DOMAIN 326 361
FT	DOMAIN 362 518
FT	DOMAIN 519 564
FT	DOMAIN 565 720
FT	DOMAIN 721 800
FT	DOMAIN 801 1331
FT	ACT SITE 162 163
FT	ACT SITE 257 257
FT	NUCLEOPHILE (BY SIMILARITY).
FT	POTENTIAL:
FT	BETA-MANNANASE/ENDOGLUCANASE A.
FT	CATALYTIC (MANNAASE ACTIVITY).
FT	PRO/SER/THR-RICH (PT BOX).
FT	SUBSTRATE-BINDING (POTENTIAL).
FT	PRO/SER/THR-RICH (PT BOX).
FT	SUBSTRATE-BINDING (POTENTIAL).
FT	PRO/SER/THR-RICH (PT BOX).
FT	CATALYTIC (MANNAASE ACTIVITY).
FT	PROTON DONOR (BY SIMILARITY).

[5] VARIANTS ADPKD P-2993, R-3016 AND V-3511, AND VARIANTS M-3510 AND P-4190. Hum. Genet. 60:119-124(1991).

[6] MEDLINE: 97342914; Pubmed:9195561; Peralta B., Gamble V., Strong C., Ong A.C.M., Sloane-Stanley J., Zerai B., Winerals C.G., Harris P.C.; "Identification of mutations in the duplicated region of the polycystic kidney disease 1 gene (PKD1) by a novel approach."; Am. J. Hum. Genet. 60:1139-1410(1997).

[7] MEDLINE: 97414058. VARIANT ALA-4058. Constantinides R., Xenophontos S.L., Neophytou P., Nomura S., Pierides A., Constantinou-Deltas C.B.; "New amino acid polymorphism, Ala/Val4058, in exon 45 of the polycystic kidney disease 1 gene: evolution of alleles."; Hum. Genet. 99:644-647(1997).

[8] VARIANTS T-2760; P-2761; V-2763; T-2764; Q-2791; T-2826; L-3008 AND L-3064. MEDLINE: 97449169; Pubmed:92855784; Wainick F.J., Piontek K.B., Cordal T.M., Weber H., Gandelph M.A., Alan F., Lens X.M., Neumann H.P.H., Germino G.G.; "A novel mutation in the PKD1 gene, duplication portion of PKD1 is revealed by use of a novel strategy for mutation detection."; Hum. Mol. Genet. 6:1473-1481(1997).

[9] VARIANT ADPKD THR-3678. MEDLINE: 97403939; Pubmed:9259200; Turco A.B., Rossetti S., Bresin E., English S., Corra S., Wainick F.J.; "Mutations of the PKD1 gene in Italian families with autosomal dominant polycystic kidney disease."; Hum. Mutat. 10:164-167(1997).

[10] VARIANT ADPKD ASP-4032, AND VARIANT VAL-4045. MEDLINE: 98180692; Pubmed:9521593; Daniels C., Maheshwar M.M., Lazarou L., Davies F., Coles G., "Novel and recurrent mutations in the PKD1 (polycystic kidney disease) gene."; Hum. Genet. 102:216-220(1998).

[11] VARIANT ADPKD MET-3375. MEDLINE: 99118881; Pubmed:9921908; Constantinides R., Kyriakides G., Hadjigavriel M., Kasai H., Xenophontos S., Zerai B., Winerals C.G.; "Late heterozygous mutation in the polycystic kidney disease with a missense mutation in the repeated region of PKD1."; Hum. Genet. 103:709-717(1998).

[12] VARIANTS ADPKD L-324 AND S-845, AND VARIANTS R-1399 AND L-1786. MEDLINE: 99293480; Pubmed:10364515; Thomas R.L., McConnell R., Whittecker J., Kirkpatrick P., Bradley J., Sweeney J.; "Identification of mutations in the repeated part of the autosomal dominant polycystic kidney disease type 1 gene, PKD1, by long-range PCR."; Am. J. Hum. Genet. 65:39-49(1999).

[13] VARIANTS ADPKD I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[14] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[15] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[16] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[17] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[18] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[19] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[20] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[21] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[22] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[23] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[24] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[25] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[26] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[27] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-1571(1999).

[28] VARIANTS ADPKD 3994 I-1-F-3906 DUPL., G-4136 AND C-4154, AND VARIANTS R-2533 AND R-2548. MEDLINE: 20046890; Pubmed:10577909; Wainick T., Phakdeekitcharoen B., Johnson A., Gandelph M., Wang M., Briefer G., Klinger K.W., Kimberling W., Gabow P., Germino G.G.; "Mutation detection of PKD1 identifies a novel mutation common to three families with end-stage renal insufficiency and/or very-early-onset disease."; Am. J. Hum. Genet. 65:1561-157

of 146 unrelated patients.*;
 RL Hum. Genet. 105:231-239(1999).
 RP [14] VARIANTS ADPKD 3748-R-V-3752 DEL AND L-4132 DEL, AND VARIANT V-4045.
 RP MEDLINE-20112427; PubMed-10647501;
 EX Afzal A.R., Hand M., Ternes-Pereira E., Saggart-Walik A., Taylor R.,
 RA Jeffery S.;
 RA "Novel mutations in the 3 region of the polycystic kidney disease 1
 (PKD1) gene.*";
 RL Hum. Genet. 105:648-653(1999).
 RP [15] VARIANTS ADPKD PRO-4225 AND TRP-4276.
 RP MEDLINE-99217041; PubMed-10200984;
 EX Badenas C., Torra R., San Millan J.L., Lucero L., Mila M.,
 RA Estivill X., Darnell A.;
 RA "Mutational analysis within the 3' region of the PKD1 gene.*";
 RL Kidney Int. 55:1225-1233(1999).
 RP [16] VARIANTS ADPKD MET-2250 AND TRP-2329, AND VARIANTS CYS-2379; LEU-3066;
 RP VAL-3139 AND LEU-3193.
 RP MEDLINE-20311156; PubMed-10854095;
 EX Perrichot R., Mercier B., Quere I., Carre A., Simon P., Whebe B.,
 RA Clodes J., Ferec C.;
 RA "Novel mutations in the duplicated region of PKD1 gene.*";
 RL J. Hum. Genet. 8:353-359(2000).
 RP [17] VARIANTS ADPKD PRO-2921 AND MET-3375, AND VARIANT LEU-3066.
 RP MEDLINE-20382887; PubMed-10923040;
 EX Koptydes M., Mean R., Demetrou K., Constantinides R., Pierides A.,
 RA Harris P.C., Deltas C.C.;
 RA "Screening of the PKD1 duplicated region reveals multiple single
 nucleotide polymorphisms and a de novo mutation in Hellenic
 polycystic kidney disease families.*";
 RL Hum. Mutat. 16:176-176(2000).
 RP [18] VARIANTS ADPKD GLN-3719 AND PRO-3852, AND VARIANT VAL-4045.
 RP MEDLINE-20514565; PubMed-11058904;
 EX Aguiari G., Savelli S., Garbo M., Bozza A., Augello G., Penciazzi L.,
 RA Peollicelli E., La Torre C., Cappelli G., Piva R., del Senno L.;
 RA "Novel mutations in the duplicated region of PKD1 gene in Italian
 polycystic kidney disease 1 (PKD1) gene: expression of mutated
 genes.*";
 RL Hum. Mutat. 16:444-445(2000).
 RP [19] VARIANTS ADPKD SER-1166; GLU-1956; CYS-2408 AND GLY-2442--2443 INS,
 RP AND VARIANTS HIS-1995 AND ASN-2604.
 RP MEDLINE-20275368; PubMed-11012873;
 EX Phakdeekitcharoen B., Watnick T.J., Ahn C., Whang D.-Y., Burkhardt B.,
 RA Germino G.G.;
 RA "Thirteen novel mutations of the replicated region of PKD1 in an Asian
 population.*";
 RL Kidney Int. 58:1400-1412(2000).
 RP [20] VARIANTS ADPKD TRP-3753 AND ASN-3815.
 RP MEDLINE-20275368; PubMed-10729710;
 EX Kim U.K., Jin D.K., Ahn C., Shin J.H., Lee K.B., Kim S.H., Chae J.J.,
 RA Hwang D.Y., Lee J.G., Namkoong Y., Lee C.C.;
 RA "Novel mutations of the PKD1 gene in Korean patients with autosomal
 dominant polycystic kidney disease.*";
 RL Mutat. Res. 432:39-45(2000).
 RP [21] VARIANTS ADPKD Q-13; R-75; G-139; 1992-F-T-1993 DELINS L; 2220-R-P-
 RP 2224 DEL; D-2336; D-2752; 2762-11MR-2765 DUPL; N-2768; K-2771; P-2816;
 RP S-2858; 3012-T-Y-3017 DEL AND 3748-L-R-3752 DEL, AND VARIANTS S-
 RP 2674; N-2708; T-2734; L-2735; C-2765; M-2782; R-2814; G-2888; I-2905;
 RP D-2966 AND L-3066.
 RP MEDLINE-21063179; PubMed-11115377;
 EX Rossetti S., Strmecki L., Gamble V., Burton S., Sneddon V., Peral B.,
 RA Roy S., Bakaloglu A., Komal R., Winealis C.G., Harris P.C.;
 RA "Mutational analysis of the entire PKD1 gene: genetic and diagnostic
 implications.*";
 RL Am. J. Hum. Genet. 68:46-63(2001).
 RP [22]

Query Match 3.78; Score 112.5; DB 1; Length 4303;
 Best Local Similarity 18.34; Pred. No. 12;
 Matches 100; Conservative 64; Mismatches 178; Indels 203; Gaps 22;
 QY 77 RVOALTSDSPALVGSNITFVNLVPRCKEDANGNIYVERNCRSDRLASDPVYVMT 136
 DB 1018 RMGQLQVSTVPAPLSPNATLALTA-----GVLV-----DSAVEA-----FLMT 1056
 QY 137 TGADDEMDENTSGQLRFPDCKPFPBPHGRKKNFVYVHTLQVFKQKGGCARSV1 196
 DB 1057 RS-----DGFALHGFQF-----PYNSFPVP----- 1078
 QY 197 NTVNLTVGPQWMEIV-----FRHCHGAPIPSKVQDVIVTDQIPFIYVMYKNDNRN 249
 DB 1079 -----DPSVAQVIVHEHNVHTFYAAFGCEVLLTVLASNAFENITQQVPVSV-----RA 1124
 QY 250 S-----SDTVLRDLPIFPOVLHHDPSHLNYSALSYKNNFQNTGLGVSNHITLNI 301
 DB 1125 SLPSVAVCVSDGLVAVAGRVTF-----YPHPLSPGGLVITMDFGSGPVLTSQSPAAHH 1179
 QY 302 TVLNGTTFENITVQTAVGFCP-----SPTSPSSSTSPS 337
 DB 1180 TYASRQTYHVRLENNVNTSGAAQADVRVFEELAGLSVDMSLAVEGQANV---VWSAAVQ 1236
 QY 338 PASGSPSTLSTFSLSMPTGYKSMELSDSNENCR-----NRKCYFRATITVIGLLEV 392
 DB 1237 TCDNITMTFDMGDTGLVSGPEATVEHVLRAGNCTVTVGAGSPAGHLASLHVLVFELEV 1296
 QY 393 NIQVADVPITLQPDNSIMDF-----VTCGKATPTEACTIISDP 433
 DB 1297 LRVEPA-ACIPT-QDARLTATYGNFAHYLPDMTFDGSNSTVVRGC-FVTHNPTFSG 1353
 QY 434 TCOIA-----ONR-----VCSFVAV-----DELCLLS----- 455
 DB 1354 TFLALVLSRRNHAHYFYSICVPEYGVNVTLOPERQVFGDEANLVACAMPFFPYRYT 1413
 QY 456 -----VRRFNGSGTYCVNFTLGGDASLAUTSALISIPCKDLGSLPRT 498
 DB 1414 WQFQTEEAATPANGPEVFFIRDPGSLVTVVYTSNNLSAANDSALVEQEPVLVTSIK- 1472
 QY 499 VNGVL 503
 DB 1473 VNGSL 1477
 Search completed: April 7, 2003, 16:08:40
 Job time : 23 secs

Result	Query No.	Score	Query Match	Length	DB	ID	Description
1	3061	100.0	572	11	09QZF6	09QZF6	rattus norv
2	2706	88.4	574	11	09QXA0	09QXA0	mus musculus
3	2700	88.2	574	11	09Q9P1	09Q9P1	mus musculus
4	711	23.2	206	4	09GF58	09GF58	homo sapien
5	634.5	20.7	626	11	09C2B2	09C2B2	mus musculus
6	398	13.0	461	6	097884	097884	equus caball
7	268	8.8	270	13	093391	093391	coturnix co
8	121	5.2	136	11	09Q160	09Q160	mus musculus
9	121	5.2	136	11	09Q160	09Q160	mus musculus
10	146	4.8	1598	17	09R9K7	09R9K7	mus musculus
11	139.5	4.5	1673	1	0977W5	0977W5	methanosarc
12	137.5	4.5	9931	10	09AC44	09AC44	chlamydomon
13	134.5	4.4	955	10	09AF92	09AF92	chlamydomon
14	132	4.3	936	2	09AQH0	09AQH0	caldicellul
15	130.5	4.3	688	17	09TR88	09TR88	methanosarc
16	128.5	4.2	34	12	091GH4	091GH4	epiphyas po

[illegible]

```

Query Match      88.2%; Score 2700; DB 11; Length 574;
Best Local Similarity 88.8%; Pred. No. 2e-220;
Matches 507; Conservative 20; Mismatches 42; Indels 2; Gaps 1;

Oy 1 MESICGVLFLLAAGLPAAKRFVDVGHGQYFQWNRNQLRGWSSDENMDKLYP 60
Db 1 MESICGVLFLLAAGLPAAKRFVDVGHGQYFQWNRNQLRGWSSDENMDKLYP 60
Oy 1 MESICGVLFLLAAGLPAAKRFVDVGHGQYFQWNRNQLRGWSSDENMDKLYP 60
Db 1 MESICGVLFLLAAGLPAAKRFVDVGHGQYFQWNRNQLRGWSSDENMDKLYP 60
Oy 61 VNRGEGRKKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGNTVYERNK 120
Db 61 VNRGEGRKKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGNTVYERNK 120
Oy 121 RSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYVFTL 180
Db 121 RSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYVFTL 180
Oy 61 VNRGEGRKKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGNTVYERNK 120
Db 61 VNRGEGRKKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGNTVYERNK 120
Oy 121 RSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYVFTL 180
Db 121 RSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYVFTL 180
Oy 181 GQYFQ 185
Db 181 GMLQ 185
Oy 181 GQYFQ 185
Db 181 GMLQ 185

RESULT 5
OQCZB2 PRELIMINARY; PRT: 626 AA.
ID OQCZB2 2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DN N/A.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RA [1] TaxID=10090;
PC SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=EMBRYO;
MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Saito M., Iwata M., Nishikawa T., Ohtsuka T., Yamashita M.,
RA Kadoya K., Matsuda H.A. Ashburner M. Baralov S. Casavant T.,
RA Fleischmann W., Gaasterland T., Glasl C., King B., Kishida H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Paoletti G., Quackenbush J.,
RA Schriml J.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anoh H., Beldarelli R., Barsh G.,
RA Akaike Y., Iwata M., Nishikawa T., Ohtsuka T., Yamashita M.,
RA Bastein M., Bult C., Pridmore D., Rhee S.Y., Rhee S.Y.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welt C., Whittaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012808; BAB28486.1; -.
DR MGI; MGI:96301; S1.
DR InterPro; IPR000601; PKD_domain.
DR SMART; SMART:SM00083; PKD_1.
DR PROSITE; PS00093; PKD: 1.
SQ SEQUENCE 626 AA; 66301 MW; 78C0A06C53212674 CRC64;

Query Match      20.7%; Score 634.5; DB 11; Length 626;
Best Local Similarity 27.9%; Pred. No. 3.7e-45;
Matches 177; Conservative 88; Mismatches 191; Indels 179; Gaps 20;

Oy 54 WDEGLVYVRRGEGKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGN 113
Db 54 WDEGLVYVRRGEGKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGN 113
Oy 43 WNRCLVPEVTEVOG--SNCWRGGVSLRVINDGPTLVGANASFIALHFGSKVLPDQ 100
Db 43 WNRCLVPEVTEVOG--SNCWRGGVSLRVINDGPTLVGANASFIALHFGSKVLPDQ 100
Oy 114 IYVENCNRSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYV 157
Db 114 IYVENCNRSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYV 157
Oy 101 VI-----WANTLINGSWGQVYVFPQEDDCAFP 132
Db 101 VI-----WANTLINGSWGQVYVFPQEDDCAFP 132
Oy 158 DCKEFPFRPHGKKNFYVFTLQYFQKLGCCSARVINTVNLVGVQVMEVTFRRHG 217
Db 158 DCKEFPFRPHGKKNFYVFTLQYFQKLGCCSARVINTVNLVGVQVMEVTFRRHG 217

```

```

Db 1 MBLVYFLGFLLLAALPLDAKRFVDVGHGQYFQWNRNQLRGWSSDENMDKLYP 60
Oy 61 VNRGEGRKKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGNTVYERNK 120
Db 61 VNRGEGRKKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGNTVYERNK 120
Oy 121 RSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYVFTL 180
Db 121 RNEAGLSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYVFTL 180
Oy 181 GQYFQ 185
Db 181 GMLQ 185

RESULT 5
OQCZB2 PRELIMINARY; PRT: 626 AA.
ID OQCZB2 2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DN N/A.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RA [1] TaxID=10090;
PC SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=EMBRYO;
MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Saito M., Iwata M., Nishikawa T., Ohtsuka T., Yamashita M.,
RA Kadoya K., Matsuda H.A. Ashburner M. Baralov S. Casavant T.,
RA Fleischmann W., Gaasterland T., Glasl C., King B., Kishida H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Paoletti G., Quackenbush J.,
RA Schriml J.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anoh H., Beldarelli R., Barsh G.,
RA Akaike Y., Iwata M., Nishikawa T., Ohtsuka T., Yamashita M.,
RA Bastein M., Bult C., Pridmore D., Rhee S.Y., Rhee S.Y.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welt C., Whittaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012808; BAB28486.1; -.
DR MGI; MGI:96301; S1.
DR InterPro; IPR000601; PKD_domain.
DR SMART; SMART:SM00083; PKD_1.
DR PROSITE; PS00093; PKD: 1.
SQ SEQUENCE 626 AA; 66301 MW; 78C0A06C53212674 CRC64;

Query Match      20.7%; Score 634.5; DB 11; Length 626;
Best Local Similarity 27.9%; Pred. No. 3.7e-45;
Matches 177; Conservative 88; Mismatches 191; Indels 179; Gaps 20;

Oy 54 WDEGLVYVRRGEGKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGN 113
Db 54 WDEGLVYVRRGEGKMSWEGGRVQAALTSDSPALVGSNTFVNLVFPKCKEDANGN 113
Oy 43 WNRCLVPEVTEVOG--SNCWRGGVSLRVINDGPTLVGANASFIALHFGSKVLPDQ 100
Db 43 WNRCLVPEVTEVOG--SNCWRGGVSLRVINDGPTLVGANASFIALHFGSKVLPDQ 100
Oy 114 IYVENCNRSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYV 157
Db 114 IYVENCNRSDELSADPYVYNTGADDEWEDNTSOGHLRFPOCKFPRPHGKKNFYV 157
Oy 101 VI-----WANTLINGSWGQVYVFPQEDDCAFP 132
Db 101 VI-----WANTLINGSWGQVYVFPQEDDCAFP 132
Oy 158 DCKEFPFRPHGKKNFYVFTLQYFQKLGCCSARVINTVNLVGVQVMEVTFRRHG 217
Db 158 DCKEFPFRPHGKKNFYVFTLQYFQKLGCCSARVINTVNLVGVQVMEVTFRRHG 217

```

```

Db 133 DGGCFSGKPKPRGFTYVWKWKWVLOGPVSRSLSTNGHMLGTHHEVTVYHRG 192
Oy 218 -RAYIPISKVDDVYVTDLPPIFYWYQKNDRNSDETFELDPIFPDVLIDHPSHLNY 276
Db 193 SOSTVPLAHSSFFITDQVPVSVSQAQDDEGTFKRLNHLPLFALQALHDHDSGLAE 252
Oy 277 SAISKWKNFGTGLFVSNHHTLHNYVLTGTFNLTVOYAVP-----GPCPS----- 326
Db 253 ADLSTWMDGDTGLTSLRAIDVTHITLDESGVTAQVVLQAAIPLVSGSSVPPTIDGY 312
Oy 327 -----TPSPSSPSS-----FASPSPPLS----- 347
Db 313 MPTAEPGTTSGROCTTKVVGCTPGQMPPTQISCTTVVQMPTEVTATTEQOMLSAVID 372
Oy 348 -----TPS-PS-----LMPYGYKSKNELSDISNE----- 369
Db 373 TILAIVSTEGCTGTPRGSGTVAQNTTEGPDASPLPFGSSGTSPLLDITDTML 422
Oy 370 -----NCRINRYGVFRATITVXGLLEWNIQVADVPIPTLQPNLSMDIFVTCCKGAT 432
Db 433 VKROVPLDCLVLYRGFSALADIVOGIESAEILQA--VPF-----SEGDAFELTVSOGGL 486
Oy 423 PTEACTIISDPTQIAQNRVCSPVAVDELCLLSVRRAF-NGSGYCVNFTPLDGDASIALT 481
Db 487 PRACNDLSSGCGCPAQRQGLSVSPSPDQVLYLQVLGSGSTGLVSLADANSUVA 546
Oy 482 SALISTPGKDJG---SPIRTVNGVLISIGCLAMFYVMTILLYKKHYKTYKPTGNCTRN 538
Db 547 STQLVVPQDGGIQAQPL-----LVGILLVAVVLASLHRIIR-----L 586
Oy 539 KQKGLSVFLSHAKAF-----SKDREKDPILL 565
Db 587 KQGSVSVQAPRGSTHKLRLPPVFRAGLGENSPILL 621

RESULT 6
Oy97884
ID O97884 PRELIMINARY; PRT: 461 AA.
AC O97884
RT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE Melanocyte protein 17 (Fragment).
GN PMEL17.
OC Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Perissodactyla; Equidae; Equus.
RN [1] NCBI_TaxID=9786;
RN [1] PubMed=12976;
RC TISSUE=SKIN;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
RA "An equine sequence homologous to melanocyte protein 17 (PMEL17)
RA mapped to chromosome 6p23." In: EMBL/GenBank/DBJ databases.
DB PMEL1 AF076780; AAC37108.1.
DR InterPro: IPR000601; PKD_domain.
DR Pfam: PF00801; PKD; 1.
DR SMART: SMO089; PKD; 1.
DR PROSITE: PS0093; PKD; 1.
FT NON_TER 461
FT SEQUENCE 461 AA: 49334 MW: 12752AF6C1EC373D CRC64;

Query Match
Best Local Similarity 13.0%; Score 398; DB 6; Length 461;
Matches 108; Conservative 65; Mismatches 122; Indels 144; Gaps 12;

Oy 215 RGRATYPISKVDVYVTDLPPIFYWYQKNDRNSDETFELDPIFPDVLIDHPSHL 274
Db 1 RGSQTYWPLAUSNSAPITDQVPVSVSQAQDDEGTFKRLNHLPLFALQALHDHPSGL 60

```

```

Oy 275 NYSAISYKRWPDNTGLFVSNHHTLHNYVLTGTFNLTVOYAVP-----GPCPS----- 325
Db 61 AGADLSYTMFGDGTGLTSLRALVTHITLDESGVTAQVVLQAAIPLTSCGSSVPVGTG 120
Oy 326 -----TPSPSSP-----FASPSPPLS----- 344
Db 121 GYVFAEAPGTTAGQVPTADVNTTPQGVPTAEPSRTTAVQVPTTEVISTPTVOVPTAED 180
Oy 345 TLTSP-----SPSLMPTGYKSKNELS-----DISNE----- 369
Db 181 LQTTPEQVTPESLGLTTLAEMPLQRLGIPPEVSNRSLGHGNSQVTVSMWKPMQRCTQ 240
Oy 370 -----NCRINRYGVFRATITVXGLLEWNIQVADVPIPTLQPNLSMDIFVTCCKGAT 432
Db 241 PLUSLRVQPAHSCRRNRYRSOSPLDGTATFLVKRQVPLDCLVLYRGFSALADANSUVA 449
Oy 390 LEVNIQVADVPIPTLQPNLSMDIFVTCCKGATPTEACTIISDPTQIAQNRVCSPVAVD 449
Db 301 ESAEILQA-----VPSSRGD--AFELTVSOGGLPKACMDLSSGCGCPAQRQGLQVPTPN 354
Oy 450 ELCLLSVRRAF-NGSGYCVNFTPLDGDASIALTSIPGKDJG---SPIRTVNGVLIS 505
Db 355 PAQQLVTHQVLYKSGSGYCVNLSADANSUVAQVSTQVLMVPGQAGLQAPL-----F 406
Oy 506 IGLAMFYVTR-TILLYKK 523
Db 407 VGILLVLAANVLASLTVRR 425

RESULT 7
Oy93391
ID O93391 PRELIMINARY; PRT: 270 AA.
AC O93391
RT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE Melanosome matrix protein (Fragment).
GN MPM15.
OC Coturnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Neophala; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99119493; PubMed=9918878;
RA Oberst C., Hartl M., Welschrichen R., Bister K.;
RA "Cloning of a cell surface protein from coxycycline-controlled expression
RA of a melanocyte protein." In: EMBL/GenBank/DBJ databases.
RL Cytology 253:193-207 (1999).
DR EMBL: AF077328; AAD12180.1; -.
KW Matrix protein.
FT NON_TER 270 AA: 25785 MW: 554AA6F0B875B63 CRC64;
FT SEQUENCE 270 AA: 25785 MW: 554AA6F0B875B63 CRC64;

Query Match
Best Local Similarity 29.5%; Pred. No. 1.4e-14;
Matches 79; Conservative 36; Mismatches 111; Indels 42; Gaps 9;

Oy 318 AVGPCPSPT-----TPSSSTSPSPASSPPTL--STPSPSLMPTGYKSKNELSD 365
Db 22 ATAGATADTGTAGTAGIADPTAGTAPRFLPVLKQAPASEPFG----- 73
Oy 366 ISNENCRINRYGVFRATITVXGLLEWNIQVADVPIPTLQPNLSMDIFVTCCKGATPE 425
Db 74 ----CVLYRGVGSFTELIVQGIENVAIQV--VPAAPGSGNS-VELTVTCGSLPEE 125
Oy 426 ACTIISDPTQIAQNRVCSPVAVDELCLLSVRRAFNSGTYCVNFTLDDGASIALTSALI 485
Db 126 VCTVYVADACRQAQWCTGCVNAPPCQQLVQDQSGGLCLVSLANGCLAVSTRV 185
Oy 486 STPCQDGLSGPLTVNGVLISIGCLAMFYVMTILLYKKHYKTYKPTGNCTRNRYK----- 539

```

```

Db 186 AVGG---ASP---AAGGTTTGVGLLIVAGLTAATYTRVRYKVSPLLPFAPOVSRPHSWLEP 240
OY 540 -GRGLSVFLSHAKAPSGRGDRKDPQLQ 566
Db 241 PGAALRLLLRQAFGAPSG--ESSPLLR 266

RESULT 8
OYQV67 ID OYQV67 PRELIMINARY; PRT; 236 AA.
AC OYQV67
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SCF-matrix melanosomal protein silver (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20063689; PubMed=10594241;
RA Solano F., Esparza M., Jimenez-Cervantes C., Bennet D.C., Lozano J.A.,
RA Martinez-Esparza M., Garcia-Borrn J.C.;
RT "The mouse silver locus encodes a single transcript truncated by the
RT silver mutation.";
RL Mamm. Genome 10:1168-1171(1999).
DR EMBL; AF119092; AAF19761.1; -.
FR NON_TER 1
FT VARIANT 117 117
SQ SEQUENCE 236 AA; 24293 MW; 3ECB8124D4F10E7 CRC64;

Query Match 8.2%; Score 251; DB 11; Length 236;
Best Local Similarity 30.5%; Pred. No. 3, 26-13;
Matches 72; Conservative 32; Mismatches 104; Indels 28; Gaps 7;

OY 306 NOTFNFMLTVOTAVFGCPSPSPSSSTSPSPSSPTLSTPSPSLMPTGVKSMELSD 365
Db 3 SAVINTTLLAEVSTTEGTGTPTP-RPSGTTVAQAATITGSPASPLLTQSGTSGISPLDD 61
OY 366 ISN-----ENCKRNRYGVRATIRIVGILEVITVAVADVPITPLOFNSLMDFTV 416
Db 62 TDTIMLVKQVPDLCVLYRGSESLADIVQIESAELLQA---VPSSSD--AFELTV 115
OY 417 TCKGATPTEACTITSDPTCOIAORVCSFPAVDELCLLSVRRAP-NGSGTYCVNFTLGGD 475
Db 116 SCQGLPFPKACMDISSPGCPQAPORLCQVSPSPDQQLVHQVLKGGSGTYCLNVLASDA 175
OY 476 ASIALATSIIRGKDLG---SPILRTVNCVLISIGCLAMFVMTWTLILYKHKHTYK 528
Db 176 NSLAVASTQLVFPQDGLGQAPL-----LVGILLVAVVAVLASLIRHR 223

RESULT 9
OYQV70 OYQV70 PRELIMINARY; PRT; 141 AA.
AC OYQV70
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SCF-matrix melanosomal protein silver (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20063689; PubMed=10594241;
RA Solano F., Esparza M., Jimenez-Cervantes C., Bennet D.C., Lozano J.A.,
RA Martinez-Esparza M., Garcia-Borrn J.C.;
RT "The mouse silver locus encodes a single transcript truncated by the

```

```

RT silver mutation.*;
RL Mamm. Genome 10:1168-1171(1999).
DR EMBL; AF118450; AAF19760.1; -.
FR NON_TER 1
FT VARIANT 117 117
SQ SEQUENCE 141 AA; 14750 MW; 7E9B70E43D42A3E8 CRC64;

Query Match 5.4%; Score 164; DB 11; Length 141;
Best Local Similarity 31.0%; Pred. No. 3, 8e-06;
Matches 43; Conservative 20; Mismatches 55; Indels 32; Gaps 5;

OY 423 PTEACTITSDPTCOIAORVCSFPAVDELCLLSVRRAP-NGSGTYCVNFTLGGASLALT 481
Db 116 SCQGLPFPKACMDISSPGCPQAPORLCQVSPSPDQQLVHQVLKGGSGTYCLNVLASDA 61
OY 482 SALISIRGKDLG---SPILRTVNCVLISIGCLAMFVMTWTLILYKHKHTYKPGNCTRRV 538
Db 62 STQLVFPQDGLGQAPL-----LVGILLVAVVAVLASLIRHR-----L 101
OY 539 KKGKGLSVFLSHAKAPF-----SRGDRKDPILL 565
Db 102 KKGKSVSQMPGHSGTHMLRPVFRARGLGENSPILL 136

RESULT 10
OYQV71 ID OYQV71 PRELIMINARY; PRT; 1698 AA.
AC OYQV71
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein MA1904.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OX NCBI_TaxID=2214;
RN SEQUENCE FROM N.A.
RA MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Salinov S., Atnoor D., Brown A.,
RA Li N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Zimmer A., Barber R.D., McGeehan K., Talamas J., Tirrell A., Ye W.,
RA Leight J.R., Ingram-Smith C., Kustiner H.C., Grady D.A., Guss A.M.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reese J.N., Smith K.,
RA Ferry J.C., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA and physiological characterization of Methanosarcina acetivorans
RA and physiological characterization of Methanosarcina acetivorans
RL genome Res. 12:532-542(2003);
DR EMBL; AE010872; AAM05308.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1698 AA; 178126 MW; 395E1DA009A846E CRC64;

Query Match 4.8%; Score 146; DB 17; Length 1698;
Best Local Similarity 21.9%; Pred. No. 0.0042;
Matches 11; Conservative 71; Mismatches 178; Indels 158; Gaps 29;

OY 158 DCKPFP--PRPHCKKNFVYVHTLQYQKLGQC-SARVSIINTVLTVGPQVMEVTPR 214
Db 265 EGAPFGVAVTPDGAK----VYVTH-VQYFSTVAVIDVATNKVTATIPGPDGPVAVTP 319
OY 215 RHGRAVPIKSRVDYVVI-----TDQIPF-----FW-----TWQK---NDRN 249
Db 320 DGTKFTVAINCWYTSVIDTATNTVATMPVKGSPRASGRFGTSVPQVPTFSANFNMI 379
OY 250 SSDETFLDLPITFDVLIDHPSHFLYSAISYKNWFCNDTGLFVSNHNTLHNTVYLVNGTF 309

```

```

Db 380 TSDYVEL...SVQFT-----DLSENATKNNMDFGDS--FTSIKONPVHTVSEVGT 477
Qy 310 NFNLIQTAVPQCFSPPTSPSSSTSPASSPPT-----LSTSPSLMPTGYKSMEL 363
Db 428 TVKLTVSNN-----NGTDSKLTIVVNVKGLAPSTAIANNSTVSVINTGSLTT 481
Qy 364 S-----DISNENCHI--NRKYGFRTATIVDGLLE--VNIQVADVTP--TL 405
Db 482 TVPWGPGGLVAASPDGTITVNTSEYNYGTGVSVIDTALNEVATVQIDCKYSKGIATV 541
Qy 406 OPD-----NSLMDEFTVTKGATPTPEACTI-----ISDPTCQIAONRVCSPTVA 447
Db 542 TPQCKKLVANMRDIDGVSVIDTINTVIADIPVGINFLGVAAMPDGTGVVTVNNVSNVS 601
Qy 448 VDEI-----CLLSVRRAFGSGTY--CVNPTLG--DDASLALSALSISI 487
Db 602 VIDATNKYVATVTKSGPCPTVW--EGTNLVANCENWTLIIDGDS--NNTASV 656
Qy 488 FQKDLGSG-PLRTVGVLIISGLANFV-----TWVTILLYKKHK----- 525
Db 657 PA---GTWPM---GVAVSPDGTGVVYVANERNSNVSYDILATKTDIAVAVKRGCPGIATV 709
Qy 526 -----TKPDIGNCTNNVKKGKLSVF-----LSHAKAPFS 555
Db 710 TPDCTRVYVANCQNLQKRGYSIIDPATNKVIATVTKGFS 750

RESULT 11
Q977V5
ID Q977V5 PRELIMINARY; PRT; 1673 AA.
AC Q977V5;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 21, Last annotation update)
DE Surface antigen.
OS Methanosarcina mazel (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RA Takagi J, Springer T.A.;
RT "Corrected sequence of surface antigen from Methanosarcina mazel S-6";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF34229; AAK84029.1;
DR InterPro; IPR001258; NHL.
DR InterPro; IPR000501; PKD_domain.
DR Pfam; PF00301; PKD_1.
DR Pfam; PF00301; PKD_1.
DR PROSITE; PS50093; PKD; 10.
SQ SEQUENCE 1673 AA; 175096 MW; 36E8C803AA7F11E CRC64;

Query Match 4.6%; Score 139.5; DB 1; Length 1673;
Best Local Similarity 20.9%; Pred. No. 0.014;
Matches 103; Conservative 73; Mismatches 199; Indels 117; Gaps 25;

Qy 165 PBGRKWNFYVHTLQGVFKQGLQCSARVINTVNLTVGQWVEVIFRRHGRAYIPIS 224
Db 251 PBSTK-----AVTVN-VDKYFNTVSMIDGTGNKITARIPVGPDPAGIATVPGKVKYVYALS 305
Qy 225 KVKDYVVI-TDQIPFVTVYQKNDRSSDTEFLRDLPT-----FFD-----VLHID 269
Db 306 FQSPVSDTATNTATITAMVAGKNPYASQ-QFGSIPQVQVPSADFKNSTSGTIFLSE 364
Qy 270 PSFPR--LAYSALSYKNGFGDGLVGLFNNHNTLNHRTVYLVNCTFNFNLTVQAVPGCPSP 327
Db 365 PVQFTDLKSDKATDEWSPFGDSSS---SKQNPHTYSETGIYTVRLTVSNN-----NCT 415
Qy 328 PPSASSTSPSPASSPPT-----LSTSPSLMPTGYKSMELS-----DISNENC 371

```

```

Db 416 DSOISTVNVVLKSGPTSPSYATIGLSNTVSVFNTGNTLTAKTVFGNDPMGVAISPDT 475
Qy 372 RI-----NRKYGFRTATIVDGL--EVNIQVADVTP-----TLQPDN----- 409
Db 476 RYVNTNTGT-AGSVSYVDTARGEVITIVQGNKYSCGIAVTFDCKLYSDRDINGV 534
Qy 410 SLMDFTVTKGNPTE-----ACTILISDPTCQIAONRVCSPVAV-----DELCLLSVRR 458
Db 535 SVLDTSTNTVATVPAGINPLGVATIPDGRKAVYVANNYSNNVSDIVTNVNEAAVYKGT 594
Qy 459 -----AFNSGTS--YCVN-----FTLGADASIALTSAL-----ISTPGKDLGSLPRTVN 500
Db 595 GPGCVSPNDSTRLVYTCESNVSVIDTATNTVDTLAVKWPGLQVSVSPDGTGKIYVAN 654
Qy 501 GVLIISGLAMPVTVVTLILYKKHKTY--KP-----IGNCTNNVKKGKLSVF----- 546
Db 655 ERSNNVSVIDARTKNTVATAIKVGRSPGIATVDTGTVKYVANCNNENLGTISIIDAT 714
Qy 547 -----LSHAKAPFS 555
Db 715 NRVYATEKAGFS 726

RESULT 12
Q94C44
ID Q94C44 PRELIMINARY; PRT; 991 AA.
AC Q94C44;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hydroxyproline-rich glycoprotein VSP4.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-93372571; PubMed-7689882;
RA Waffenschmidt S., Woessner J.P., Bear K., Goodenough U.W.;
RT "Isolation of a cDNA encoding a stress-linking mediator involved in the inactivation of cell walls in Chlamydomonas";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF34229; AAK83527.1;
DR InterPro; IPR001412; TRNA-synt.1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN.1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN.1.
SQ SEQUENCE 991 AA; 104614 MW; 6B3CB8B331CD800 CRC64;

Query Match 4.5%; Score 137.5; DB 10; Length 991;
Best Local Similarity 22.8%; Pred. No. 0.01;
Matches 68; Conservative 37; Mismatches 104; Indels 89; Gaps 13;

Qy 246 NDNS-----SDPTFLRDLPTFEDVLLHDPSPFLNLSAISYKNGDNT----- 289
Db 620 NDANSADFCSAGADET-ANVTITVYDCLLVGLHLSQGVVAMHOITADNNPTLYAW 678
Qy 290 GLEFVNNHTLNHTVYLVNGTFNFNLTVQAVPGCPSPPTSPSSSTSPASSPPTLS-- 347
Db 679 GQYVACNPAGVTPPV-----VATPSPSPSPKPSFSPSPSPSPSPSPSPSPSP 729
Qy 348 -----TPSPSLMPTGYKSMELSDI-----SNENCR-INRYGY-----FRATITVDGLLEVI 394
Db 730 PSSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 789
Qy 395 IQV-----ADVPPTPTOPDNLMDFTVTKGATPT 424
Db 790 CQVDRNSNCANRNGMDVCTPSCGTNPAPSPSPSPSPSPSPSPSTGTGTCGACAPT 849

```

QN	NCBI_TaxID=80339;
RN	[...]
KP	SEQUENCE FROM N.A.
RP	STRAIN=TOK7B.1;
RX	MEDLINE=20171169; PubMed=1070665;
RA	Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
RT	Bergquist P.L.;
RL	Tetradomain and multifunctional glycosyl hydrolases from the extreme
RM	thermophilic Candidatus Caldococcus isolate TOK7B.1.*;
RE	CENL: AF078038; AAK0398.1; ...
DR	HSSP: QO5851; INBC
DR	InterPro: IPR001956; CBD_3.
DR	InterPro: IPROQ2860; GH_BNR
DR	InterPro: IPR002965; P_Rich_extensn.
DR	Pfam: PF01022; BNR; 9.
DR	Pfam: PF00942; CBM_3; 1.
DR	PRINTS: PR002317; PRICHEXTENS.
DR	PROSITE: PS001947; CBD_3; 1.
KW	Hydrolase.
FT	SEQUENCE 996 AA; 108275 MW; 3C72BED2F3C61A CRC64;
QT	SEQUENCE 996 AA; 108275 MW; 3C72BED2F3C61A CRC64;
Query Match 4.3%; Score 132; DB 2; Length 996;	
Best Local Similarity 21.1%; Pred. No. 0.03;	
Matches 90; Conservative 58; Mismatches 142; Indels 136; Gaps	
QY	32 EOTDHH-RENHQLRSGSDENE--WDEOLYPVWRGGEHRKDSWGGRVOALITSDSPA 88
DB	512 ELNIWFVRVGVNDKKWNPNIRIGFSYDGCKSHFGTGTSGEGTVAA--ADGSA 5699
QY	89 LV---CGSNTIFPVNLVPFCOKEDANGNVIERNCASDLASDPVYIKMTTGADDEW 144
DB	570 VWAFPKAGV-----CYSTD-NGNKWE--CAN--VPSEAIYS----- 603
QY	145 EDWTQQCHLPDPGPPRHGRKNFVVFHTLGQYFOKLQGCSARVSINTVNL--- 201
DB	604 -DAVNPKFYATNGR-----FYISADKRKTIESPAAGLPISGNFKYVPGI 649
QY	202 ---TGPOVKVEIVFRHH---GRAYIPISKVDKYVI-----TDQIPVFWYOK 245
DB	650 EGDJMLVGNNCK-----WHSTDGGSYFVKISGEDAASIGCPACETGYPIAYTKI 703
QY	246 ND---RNSEDTEFLADLPFFDVLIHDSPH-----FLNYSATS 280
DB	704 NGVGIFRSDDCKTWIR-----INDKHQGCANADICDPRVGRVFATNGLG 754
QY	281 YKANFDINGLFVSNHPLAHYVINGTFNPNIUTAVTPGCPSTPSSSSTSPSPAS 340
DB	755 LKGEIAYSNISSATPTSPSPPTPTATTI-STPAFTSTPTPTPTPTPTPTPTTP 813
QY	341 SPSPFLS-----TPSPSLMPGT-----YKSWELDSLNECRINKRYGFYRAT 382
DB	814 TPTPTPTVTPTPTPTPTPTPTPTCTGTCGLSKVLVYNKNETSASA-----GSIRPW 864
QY	383 ITIVDG 388
DB	865 FRIYWG 870
RESULT 15	
QSTR88	
ID	PRELIMINARY; PRT: 688 AA.
AC	QSTR88;
CD	QSTR88;
DT	01-JUN-2002 (TrMBUrel. 21, Created)
DT	01-JUN-2002 (TrMBUrel. 21, Last sequence update)
DT	01-JUN-2002 (TrMBUrel. 21, Last annotation update)
GN	MA1293 surface protein
OC	Methanosarcina acetivorans
OC	Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC	Methanosarcinaceae; Methanosarcinia.
OC	NCBI_TaxID=2214;

Search completed: April 7, 2003, 16:12:11
Job time : 102 secs

•

JOURNAL

Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Genetics (MAG) Cancer Genetics Office, National Cancer
Institute, 31 Center Drive, Room 11A05, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: mgc@nci.nih.gov
Tissue Procurement: ATCC/DCTP/DTF
DNA Library Preparation: Rubin Laboratory
DNA Sequencing: Genome Sequencing Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjeil, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Chang Lee, Lingling Li, Mark Mason, Candice McLeay, Steven
Ness, Pawan Pandey, David Prasad, Robert R. Rabinowitz, John
Schain, Duane Smalish, Michael Smith, Lorraine Spencer, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the NIH-MGC Consortium/ILM at: http://image.llnl.gov
Source: Tissue, NIH-MGC Consortium/ILM at: http://image.llnl.gov
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505404
This clone has the following problem: frame shifted.

FEATURES
source

1..2636
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:487773"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_id="NIH_MGC_49"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 750 a 386 c 384 g 716 t

ALIGNMENT SCORES:

Pred. No.: 2,61e-173 Length: 2636
Score: 2148.00 Matches: 393
Percent Similarity: 82.3% Conservative: 75
Best Local Similarity: 69.1% Mismatches: 86
Query Match: 10.1% Indels: 14
DB: 1 Gaps: 1

US-09-943-075A-2 (1-572) x BC025297 (1-2636)

QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
Db 52 ATGGAAATGCTCTACTATTCCTGCGGATTCCTGCTGCGGCAAGTATGCCACTTCAT 111
QY 21 AlaAlaLysArgPheArgAspValLeuValLeuValHisGluGlnIrrProAspHisPheArg 40
Db 112 GCCCAACAGATTCATGATGCTGGCATGGAAGACCTTCGCTACATGAGCGAG 171
QY 41 AsnAsnGlnLeuValGlyIrrPheSerAspGluAsnGlnIrrPheValValLeuValPro 60
Db 172 CACAATCAATTAATGCTGCTCTCTGATGAAATGACTGGAATGAAATGCTACCCA 231
QY 61 ValTrpArgArgGlyValGlyValGlyTrpPheAspSerTrpGluGlyValArgValGln 80
Db 232 GTGTGCAAGCGCGGAGACATCAGGTGGAATAAATCTCTCGAGGAGCGCGTGTGCGCGC 291
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 292 GTCTGCGGAGTACTCCAGCCCTCTGGGCTCAATATPACATATTCGGGTGAACCTG 351
QY 101 ValPhePheArgCysGlnValGlnValGlnValGlnValGlnValGlnValGlnValGln 120
Db 1390 AATGGGTCTGGAGCTACTGTGTGAACCTCACCCTGSGGCGATGACAAAGCTGGCTCTC 1449

Db 352 ATATTCCTAGATGCCAANAAGAGATGCCAATGGCAACATAGCTTATGACGAAAGCTGC 411
QY 121 ArgSerAspLeuLeuLeuLeuLeuSerAspProValTyrrAspThrPheThrGlyAlaAsp 140
Db 412 AGAATGAGCGTGGTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
QY 141 AspIuAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 472 GACAGTGGCGGAAATGAGTGGACCGGCAACCATACGCTCTCTGATGGGAA 531
QY 161 ProPheProArgProHisGlyValGlyValGlyTrpAspPheValTyrrValPheHisLeu 180
Db 532 CTTTTCCTCCACCCCGGAGGAGAGATGGAATTCATCTCATCTCTCCACACTT 591
QY 181 GlyGlnIrrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 592 GTCAGTATTTCCGAATATGGGAGGATTTCTAGTCAGAGTTCTTCGGAACACGCCAT 651
QY 201 LeuThrValGlyProGlnValMetClnValIleValPheArgArgHisGlyValGlnIrr 220
Db 652 GTACACACTGGGCTCAACTCATCGAAGTGAAGTCTCTACAGAAAGACATGGACGGCAT 711
QY 221 IleProIleSerLysValLysAspValTyrrValIleThrAspGlnIleProIlePheVal 240
Db 712 GTTCCCATACACAGATCAAGATGATGTAGTGGATGACATGATTCCTGTGTGTG 260
QY 241 ThrMetTyrrGlnLysAsnAspArgAsnSerSerAspGlnThrPheLeuArgAspLeuPro 771
Db 772 ACTATGTTCCAGAGAAGCATGCAATTCATCGAGAAACCTTCCTCAAGATCTCCCC 831
QY 261 IlePhePheAspValLeuLeuHisAspProSerHisPheLeuAsnTyrrSerAlaIleSer 280
Db 832 ATATGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
QY 281 TyrrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
Db 892 TACAAGTGGAGCTCGGGGATATATGCGCTGTGTGTTCCACCACCAATCATCTGAT 951
QY 301 HisThrTyrrValLeuAsnGlyThrPheAspPheAsnLeuThrValGlnThrAlaValPro 320
Db 952 CACAGTATGCTCTCATGGAGACCTTACGCTTACCTCACTGTGTAAGCTGGCAGCC 1011
QY 321 GlyProCysProSerProThrProSerProSerSerSerSerSerSerProAlaSer 340
Db 1012 GGAACCTTGTGCGCCACCCACACCCACCCACCT----- 1047
QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrrIrrSer 360
Db 1048 -----TCAAAACCCGCTCTCTTAGGAGCTGCTGTGTGACACCC 1089
QY 361 MetGluLeuSerAspIleSerAsnGluAsnGlyIleAsnArgTyrrGlyTyrrPheArg 380
Db 1090 CTGGAGTGGATGAGTCTCTGTAACAACTGCGAGATTACAGATATGCGCACTTCAA 1149
QY 381 AlaThrThrThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
Db 1150 GGCACCTCATCATTTAGAGGATTTAGAGGATTTACATCTCATCTGAGTACAGAGCTC 1209
QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
Db 1210 CTGATCGCGGCTGCATGCGCTGAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGG 1269
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
Db 1270 AGCATTCCACCGAGGTCTACATCTCTGTGACCCACCTGCGAGTACACCGAGAC 1329
QY 441 ArgValCysSerProValAlaValAspGluLeuLeuLeuSerValArgArgAlaPhe 460
Db 1330 ACAGTCTGACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
QY 461 ArgGlySerGlyThrTyrrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaPro 480
Db 1390 AATGGGTCTGGAGCTACTGTGTGAACCTCACCCTGSGGCGATGACAAAGCTGGCTCTC 1449

40	ATG	3AA	TGCTCT	CTACTATTTCTGGGATTTCTGCTCTGGCTCGCAAGATTGCCACTTGAT	99
21	Ala	Ala	Ile	Val	5A
100	GC	3C	CAAC	CAAGATTTCTATGATGTCTGGCAATGAAGAACTCTTGCTTACATACGGGAG	159
41	Asn	Asn	Gln	Arg	1G
160	CAC	TAC	ATCA	TAAATGGTGGTCTCTGATGAAGATTCGATGAGAAATGACCAATCCCA	219
61	Val	Trp	Arg	Gly	1G
220	GTG	TGG	AAC	GGGGAGACATGAGTGGAAAACTCTCGAAGGAGGAGCGCTGTCAGCGC	279
81	Ala	Leu	Thr	Ser	Asp
280	GTG	TG	AAC	CTGAGCTACACACCCCTCTGGGGCTCAATATATACATTTGCGGTGACCTG	339
101	Val	Phe	Pro	Arg	2C
340	ATA	TCC	TAG	TGCGCAAGGAGACATGCAATGGCAATAGCTATCTATGAGAAGACTGC	399
121	Arg	Ser	Asp	Glu	1A
400	AGM	ATG	AGCGTGGT	TATCTGCTGATCGGTATGTATTCAACTGAGCAGCATGGTGCAGAG	459
141	Asp	Glu	Asp	Trp	1G
460	GAC	TAG	TAC	GGGGAATATGGCACCGCCAAAGCCATCATACGCTCTCCCTGATGGGAA	519
161	Pro	Pro	Arg	Pro	1G
520	CCG	TTT	CTC	TACCAACCGCGATGGAGAGATGGAATTTCACTACGTCTTCCACACACTT	579
181	Gly	Ile	Trp	Phe	Gln
200	GGT	ACG	ATG	TTTCCAGAAATATGGCAGAGTTCCTGAGAGATTCTGTGACACAGACCAT	239
580	Leu	Trp	Arg	Pro	1G
640	GTG	CACTTGGG	CGCTCAACTCATGGAATGACTGTCTACAGACAGACATGACGGGCATAT	699	
221	Ile	Phe	Trp	Ser	Val
700	GTG	TCC	ATG	CAAGCTGAAAGATGTGATGCTGTCACATGATCATGCTGTGTGTTGG	759
241	Thr	Met	Tyr	Gln	1G
760	ACT	GAT	GTCCAG	AAGACATCGAAATTCATCCGACCAACCTCTCTCAAGATGCTCCC	819
261	Ile	Phe	Phe	Asp	Val
820	ATT	GAT	GTGATGGTCTGATTCATGATCT	849	
3	RESULT	3			
301	LOCUS	3139997			
301	DEFINITION	3139997			
301	ACCESSION	3139997			
301	VERSION	1			
301	KEYWORDS	GI:11001518			
301	SOURCE	human			
301	ORGANISM	human			
301	REFERENCE	3139997			
301	AUTHORS	3139997			
301	TITLE	3139997			
301	JOURNAL	3139997			
301	COMMENT	3139997			

Alignment Scores:	3.3e-73	Length:	880
Pred. No.:	972.00	Matches:	193
Score:	75.97%	Conservative:	41
Percent Similarity:	75.97%	Mismatches:	46
Best Local Similarity:	62.66%	Indels:	28
Query Watch:	31.75%	Gaps:	3
DB:	12		


```

Qy 387 pGlytLeuGluValAsnThrLeuGlnValaAspValProIlePheProThrLeuGlnPr 407
Db 441 GGAATCTTACAGATTTACATCCATCCAGATGAGAGCGCTCTGATGCGGTGCCGTC 500
Qy 407 oAspAsnSerLeuMetAspPheIleValThrCysGlyGlyAlaThrProThrGluAlaCy 427
Db 501 TGAAGCTCCCTAATAGACTTTGTCTGACCTGCTCCAGGAGGAGCATTCCTCCACGAGGCTGTG 560
Qy 427 shThrIleLeuSerAspProThrCysGlnIleIleValaAsnArgValCysSerProValaI 447
Db 561 TACATCATTTCTGACATCCATCCCTGCGATGATGACCTCCACGACACAGCTCTGAGCCCTGTGGA 620
Qy 447 aValAspGluLeuCysGlnValaThrAlaPheAsnGlySerGlyThrTrpCys 467
Db 621 TGTGGATGAATGTGTCTGCTGCTGCTGAGACGAACTTCAATGGGTCTGCGAGCTACTG 660
Qy 467 sValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
Db 681 TGTGACCTTACCTCCCTGGGGGATGACACAGCTGTGGTTTA 720

RESULT 6
Bg679094 822 bp mRNA linear EST 01-MAY-2001
DEFINITION 60267047F1 NC1_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4752178 5',
mRNA sequence.
ACCESSION Bg679094
VERSION 1
KEYWORDS EST 75094.1 GI:13910491
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH, nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Robert Strausberg, Ph.D.
EMAIL: cgapbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
Cloning Strategy: Lambda ZAP Express, M.M.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0609 row: m column: 11
High quality sequence score: 816.
FEATURES
Source location/Qualifiers
1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4752178"
/clone.lib="NC1_CGAP_Skn4"
/tissue.type="squamous cell carcinoma"
/lab.name="Bibi G. T. Hage-responder"
/orig.ident="NC1_CGAP_Skn4"
/submitter="NCI, NCI-DC, NCI-SPOG2; Site 1: NCI; Site 2: SALL; Cloned unidirectionally. Primer: oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NC1_CGAP Library."
BASE COUNT 232 a 186 c 204 g 200 t
ORIGIN
Alignment Scores:
Pred. No.: 6, 21e-71 Length: 822
Score: 945.00 Matches: 198
Percent Similarity: 83.69% Conservative: 38
Best Local Similarity: 70.21% Mismatches: 32
Query Match: 30.87% Indels: 15
Db: 12 Gaps: 0

Us-09-943-075a-2 (1-572) x Bg675094 (1-822)
Qy 22 AlalysArgPheAspValLeuGlyHisGluGlnTyProAspHisMetArgGluAsn 41

```

```

Db 2 GCAAAACCACTTCTGCTGGCAATGAAGAGCCTCTGCTATCATGAGGAGCAG 61
Qy 42 AsnGluLeuArgGlyThrSerAspGlnAsnGluGlnIleThrAspGluGlnProVal 61
Db 62 AATCAATTAATGGCTGGTCTCTGATGAATGACTGCGTAATGAATGACTTACCAGT 121
Qy 62 TPATGATGATGlyGluGlyArgTrpLys AspSerTrpGluGlyGlyArgValGlnAlaAl 81
Db 122 TGGAGCGGGGAGACATGATGGAGCAAACTCTCGGAGGAGCGCGCTGCGAGGCGGT 181
Qy 81 AlasThrSerSerProAlaLeuValGlySerAspIleThrPheValValAsnLeuVal 101
Db 182 CTGACCAAGTACTCACAGCCCTCTGGGCTCAATATATACATTCGCGTGAACCTGAT 241
Qy 101 lPheProArgCysGlnGlyGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCysAr 121
Db 242 ATTCCTCATAGATCCAAAGAGACATGCAATGCAATAGTCTATGAGAGCACTGCAG 300
Qy 121 g-SerAspGluLeuAlaSerAspProTyValTyTr-AsnTrpThrGlyAlaAsp 140
Db 302 AATATGACGCTGGTTATCTCTGATCGCTATGTTTACAAAGCTGCAGCAGCATGTCAG 361
Qy 141 AspAlaSerTrpGluAspAsnThrSerGln-GlyGlnHisLeuArg-PheProAsp-Gly 159
Db 362 GACAGTGACGGGAGAAATGGCACCGCCCAAGACCTATCATAGAGCTCTCCCTGATGGT 421
Qy 160 LysProPheProArgProIleHisGlyArgGlyLys-TrpAsnPhe--ValTyValPheHis 178
Db 422 AAACCTTTCTCCACACCCCGGATGCGAGAGATGGAACCTTCAGTCTAGCTCTCCAC 481
Qy 178 ThrLeuGlyGlnTyPheGlnLys-LeuGlyGlnCysSerAlaA-gValSerIleAsnTh 198
Db 482 ACATGTGCTATTCACAAAGATGGAGCAATGTCAGTTCAGAGATCTCTGAGACAG 541
Qy 198 r-ValAsnLeuThrValGlyGProGlnValMetGluValIleValPheA-gArgHisGlyA 218
Db 542 CAGCCATGTGACACTTGGGCTCAACTCATGGAAGTCACTGCTTACAGAGACATGGAC 601
Qy 218 gAlaTyTrIleProIleSerLysVal-LysAspValTyValIleThrAspGluIlePro 237
Db 602 GGGCATATGTTCCCATGCGACAAAGTGCAGAGATGCTAGCTGTACACAGATTCCT 661
Qy 238 lIlePhe-ValThrMetTyGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuAr 257
Db 662 GTGTACTGTCTATGTGTCAGAGACAGCATGCAAAATTCATCGGAGCAAACTCTCTCA 721
Qy 257 AspLeuProIlePhePheAspValIleIleIleAspProSerHisPheLeuAsnTySe 771
Db 722 AGATCTCCCATTTATGTTGATGCTCTGATTCATGATCTCTAGGCATCTCTCAATTATTC 781
Qy 772 rAlaIleSerTyTrLysTrpAsnPhe-GlyAspAsnThr 289
Db 782 TACCATTAC-TACAAAGTGAGCTTCGGGGGATTAATACT 818

RESULT 7
Bg551384 1071 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOUNT_6544813 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:554792
5', mRNA sequence.
ACCESSION Bg551384
VERSION 1
KEYWORDS EST 551384.1 GI:18788394
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Robert Strausberg, Ph.D.
EMAIL: cgapbs@mail.nih.gov

```

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia, Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
1 (bases 1 to 974)
NII-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straussberg, Ph.D.

Issue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.

```

SDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone sequencing by: Incyte Genomics, Inc.
Clone distribution: WCC clone distribution information can be
found at: http://www.imeg.org
http://image.llnl.gov/ I.M.A.G.E. Consortium/LLNL et;
Plate: ILAM9450 row: o column: 11
High quality sequence stop: 665.
Location/Qualifiers
1. .974
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGR:4164706"
/clone_lib="IMAGR:CGAP_S02"
/clone_note="resistant"
/notes="organ: salivary gland; vector: pCMV-SPORT6; Site:1:
Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo
DTT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library.*"
BASE COUNT 252 a 231 c 289 g 202 t
Alignment Scores:
Pred. No.: 2,51e-67 Length: 974
Score: 304.00 Matches: 169
Score Similarity: 79.72% Mismatches: 5
Best Local Similarity: 69.72% Mismatches: 47
Query Match: 29,538 Indels: 17
DB: 12 Gaps: 4
US-09-943-075A-2 (1-572) x BF29741 (1-974)
QY 1 MetGluserLeuCygSgiYvAlLeuVaIPheLeuLeuAlaAGIyLeuProLeuGln 20
Ddb 79 ATGGAAGTCTTCGCGGGGCTGTGGATATTTCTGCTGCGAGGACTGCCTCTCGAG 138
QY 21 AlaAlaIysArgPheAlaGspValLeuGluGHisGLuGlnIyPProAspHisMetArgIlu 40
Ddb 139 GTCGTCGAGCGNTTCGTGATGTGCTGGCGGCGTGAACAGTATCCCATCATCATGAGAGAG 198

```


17	CTGGGAGTTCTCTGCTCTGGCTGCAAGATTTGCCACTTGTATGTCGGCGCAACAGATTTCATCATG	76
28	ValLeuGlyHisGluGlnIleProAspHisMetArgGluAsnGlnIleuAspGlyTyrP	47
77	GTATCGGAGGAGTGAAGAACTCTTCTGCTTCTTCTGATGGGAGACACATCAATTAATGGCTGG	136
48	SeiSerAspGluAsnGlnIleProAspGluGlnIleuTyrProValTyrPArgGlyGlyGly	67
137	TCCTCTCTGATGAAGAAATGACTGGAATAGAAAACTCTACCCACTGTGGAAAGCGGAGACATG	196
68	ArCTTPtyAspSerTyrPGLuGlyValArgValGlnAlaLeuThrSerAspSerPro	87
197	AGGTGGGAAAACTCTGGAGAGAGGCGGTGTGAGGGGCTGTCCAGCAAGTGAATCAACA	256
88	AlaLeuValGlySerAsnLeuThrPheValValAsnLeuValPheProArgCysGlnLys	107
257	GCCTCTGTGGGGCTCAATATACACTTTGGGGGTGACCTGTATTCCTCTAGATGCCAAGG	316
108	GluAspAlaAsnGlyAsnIleValTyrGluArgAsnCysArgSerAspGluLeuAla	127
317	GAGATGCCCAATGACACTTATGTATGAGAAAGCTGCAATAGAGAGGCGGTTCATP	376
128	SerAspProTyrValTyrAsnTyrPThrGlyValAsnAspGluAspTyrGluAspAsn	147
377	GCTGATCTGATGTGTTTACACTTGCACAGCATGTGTCCAGAGGCAGCTGAGCGGGAATATGCC	436
148	ThrSerGlnGlyGlnIleuAsnTyrPheProAspGlyLysProPheArgProHisGly	167
437	ACCGGCAAAAGCCATCATACAGCTCTCCGTATGGGAACCTTTCTCTCACACCCCGGA	496
168	ArgLysLysTyrPAsnPheValTyrValPheLeuThrLeuGlyGlnTyrPheGluGlyLys	187
497	TGGAGAGCATGGAATTCATCTACTCTTCCACACACTTGTGTCAATTTCCAGAAATTG	556
188	GlyLysCysSerAlaArgValSerIleAsnThrValAsnLeuThrValGlyProIleuVal	207
557	GGAAGTATCTAGTAGAGATGTGAGTCTCCGTATGGGAACCTTTCTCTCACACCCCGGA	616
208	MetGluValIleValPheArgTyrHisGlyArgAlaTyrTyrPLeuSerLysValLys	227
617	ATGTAAGTGAATCTGTCTACAGAGACATGGACNGGCATATGTTCCTATC-GCACAGCTGAA	675
228	AspTyrValIleThrAspGlnIle	236
676	GATCTTTACTGTGTGATCAATNGAAT	702
RESULT 11		
LOCUS BH139360		
617852323P1 NCI-CGAP_Lu30 Mus musculus cDNA clone IMAGE:4013356 5'		
minla sequence.		
BH139360		
BH139360.1 GI:10978400		
house mouse.		
Mus musculus		
ORGANISM		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
NIH-MGC http://imgc.nci.nih.gov/		
1 (bases 1 to 943)		
Retrieved from NIH-MGC		
Contact: Robert Strausberg, Ph.D.		
Email: cgapbs@mail.nih.gov		
Tissue Procurement: Gilbert Smith, Ph.D.		
cDNA Library Preparation: Life Technologies, Inc.		
cDNA Sequencing by: Incyte Genomics, Inc.		
Clone distribution: NCI-CGAP clone distribution Information can be		
http://imgc.nih.gov		
Plate: LAM9255 row: m column: 05		
High quality sequence stop: 625.		

```

FEATURES
source
Location/Qualifiers
1. .943
/organism="Mus musculus"
/submitter="NCBI"
/ab.xref="taxon:10090"
/clone_lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
e-cadherin, transgenic model; NotI-1, expression driven by
WNT1 promoter; NotI-2, expression driven by NotI-1 primer; Oligo
dT. Library constructed by Life Technologies
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 221 a 227 c 290 g 205 t
ORIGIN
Alignment Scores:
Pred. No.: 943
Score: 1.16e-66
Matches: 186
Percent Similarity: 86.44%
Conservative: 18
Best Local Similarity: 78.81%
Mismatches: 27
Query Match: 29.27%
Indels: 9
Gaps: 2
DB: 12
US-09-943-075A-2 (1-572) x BF139360 (1-943)
QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20
Db 37 ATGGAAAGTCTCTGGGGGCTCGGGATTTCTGCTGGCTGCAGACTGCTCTCCAG 96
QY 21 AlaAlaLysArgPheArgValLeuGlyHisGluGlyTrpProAspHisMetArgGlu 40
Db 97 GCTGCCAAGCCATTCTGATGCTGCTGGCCAGCAGCAGTATCTGATCATCATGAGAG 156
QY 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTrpPro 60
Db 137 CACACCCAAATACGTGGTGGTGGTCTGGAGTGAATGATGGGATGAACACCTGTATCCA 216
QY 61 ValTrpArgGlyGlyGlyValGlyArgTrpValAspSerTrpGluGlyValGlyAlaGln 80
Db 217 GTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 276
QY 81 AlaLeuTrpSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 277 GTCTTGACCATGACTGACTACCCGCTCTGGGGTGGTGGTCCATATCATCGTTGTGACCTG 336
QY 101 ValPheProArgGlyGlyGlyValGlyValGlyValGlyValGlyValGlyValGlyVal 120
Db 337 GTGTTCCTCCAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
QY 121 ArgSerAspLeuGluAlaSerAspProTyrValTyrAsnTrpThrGlyAlaAsp 140
Db 397 AGGAATGTTG -GGAGTACTTCTTGACTCTGACTGATGCTACACTGAGCTGAGGAGGAGCAT 455
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnGlnHisLeuArgPheProAspGlyLys 160
Db 456 GATGCTGATCGGAGAGATGCCACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 515
QY 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180
Db 516 CCTTCTCTCCGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 574
QY 181 GlnGlyLeuTrpPheGlnLeuGlyGlnGlySerAlaArgValSerLeuAsnThrValAsn 200
Db 575 GTCATGATATTC -CAAAACTGGGTGGTGGTTCGAGCAGGGTCTATAAACAAAGTCAC 633
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArg---ArgHisGlyArgAl 219
Db 634 TGCACACTGG -CTCAGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 692
QY 219 atyrIleProIleSerLysLysAspValTyrValIleThrAsp 234
DB: 12

```

```

Db 693 CCATCTC-----CAAGTGAACCAAGTGTGTGTGTGATTACAGAT 729
RESULT 12
LOCUS 64
DEFINITION
AL550464 L1_NFL006_PL2 Homo sapiens cDNA clone CS01057YD14 5
pr, mRNA sequence.
ACCESSION
AL550464
VERSION
AL550464.1 GI:12897464
KEYWORDS
EST.
SOURCE
human
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 883)
AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization.
JOURNAL
Genomics (2001).
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY Cedex - France
Email: segref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1. .883
/organism="Homo sapiens"
/submitter="t200-9506"
/clone_lib="L1_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer, five prime end
of the cDNA was generated by random priming and the cDNA
cloned into the NotI-digested EcoRV site of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@litech.com URL :
http://fulllength.invitrogen.com
BASE COUNT 243 a 192 c 218 g 219 t 11 others
ORIGIN
Alignment Scores:
Pred. No.: 2.13e-66
Score: 892.50
Matches: 883
Percent Similarity: 99.62%
Conservative: 175
Best Local Similarity: 69.18%
Mismatches: 32
Query Match: 29.18%
Indels: 5
Gaps: 2
DB: 2
US-09-943-075A-2 (1-572) x AL550464 (1-883)
QY 1 MetCysLeuGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20
Db 65 ATGAGATGCTCTACTACTTCTCGGATTTCTGCTGGCTGCAAGATGCCACTTGAT 124
QY 21 AlaAlaLysArgPheArgValLeuGlyHisGluGlyTrpProAspHisMetArgGlu 40
Db 125 GCGCCCAAGCATTTTCATCTGCTGGCAATGAAGACCTTCTGCTACATGAGGAG 184
QY 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlyTrpPro 60
Db 185 CACATCAATTAATGCTGCTCTCTGATGAATGCTGGAATGAATAACTTACCA 244
QY 61 ValTrpArgGlyGlyValGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80
Db 245 GTGTGGAGAGGGGAGACATGAGGTGGAAACCTCTCGAAGAGGAGGCGGTGTCGACGC 304
QY 81 AlaLeuThrSerAspProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 305 GTCTCAGCAGCTACACAGCCCTCTGTGGCTCAATATCAATATTCGTGGTGAACCTG 164
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120

```

S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system—384-format
 microarray with 384 multicapillary sequencer. *Genome Res.*
 10 (11): 1751-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugihara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* 11 (2): 281-289 (2001)
 Kamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shingawa, A., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T.,
 Taniuchi, I., Hayashizaki, Y. and Muramatsu, M.
 Mapping of 1903 mouse cDNAs on mouse chromosomes. *J. Struct.*
Func. Genomics 2: pre-L72-L86 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.

Location/Qualifiers
 1. -623
 /organism="Mus musculus"
 /strain="C57BL/6J 0090"
 /db-ref="taxid:10090"
 /clone="4732452017"
 /clone_lib="RIKEN full-length enriched, 10 day neonate
 skin"
 /sex="mixed"
 /tissue_type="skin"
 /dev_stage="10 days neonate"
 /lab_host="DH10B"

Notes: Site 1: Salt; Site 2: BamHI; cDNA library was
 prepared by using thermostable thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper cDNA. The cDNA library was found to be enriched
 to 10.0 and subcloned to 80,100.0. Second
 strand cDNA was prepared with the primer adaptor of
 sequence [5' GAGAGAGATCTCGAGTAAATTAATTCCTCCGCCCCCC
 3']. cDNA was cloned into the XhoI and BamHI sites.
 Vector: a modified phuescript KS(+) after bulk excision
 from Lambda ϕ 19 t

BASE COUNT 141 a 157 c 139 t
 ORIGIN

Alignment Scores:
 Length: 4,036-66
 Pred. No.: 887.00
 Score: 163
 Percent Similarity: 92.93%
 Best Local Similarity: 88.59%
 Mismatches: 13
 Indels: 1
 Gaps: 0
 DB:

US-09-943-075A-2 (1-572) x BB613060 (1-623)

QY 1 Met(GluSerLeuGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
 Db 72 ATGGAAGCTCTCCGGGCTCTCGGAGATTTCTGCTGCTGCAGGATGCTCTCCAG 131
 QY 21 AlaAlaArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
 Db 132 GCTCCAGCGATTTGTGTGTGTGGCCATGCACAGTATTCCTCCAGTATCCGATGAGAG 191
 QY 41 AsnAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60
 Db 192 CACACACCAATTAACGTGCTGCTTCGGATGAAATGAATGGATGAACACCTGTATCCA 251
 QY 61 ValTrpArgArgGlyGluGlyValArgTrpLysAspSerTrpGluGlyArgValGlnAla 80

365 ATATTCCTCCATAGATCCCAAGAGAGATGCCAATGGCAACATGCTATGAGAGAGATCTC 424
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyLysAsp 140
 Db 425 AGAATAGAGCTGGTTATCTGCTGATGCTGCTGATGTTTACAACTGGACAGCATGGTCAGAG 484
 QY 141 AspGluAspTrpGluLysAspAsnThrSerGlnGlnGlnHisLeuArgPheProAspGlyLys 160
 Db 485 GACAGTGGAGGGGAAATGGCACGGCAAGGCATCATACAGCTCTCTCCCTGATGGAAA 544
 QY 161 ProPheProArgProPheHisGlyArgLysLysTrpAsnPheValTyrValPheHisIsthLeu 180
 Db 545 CTTTCTCTCACACCCCGGATGGAGAGATGGAAATTCATCTACGCTCTCCACACACTT 604
 QY 181 GlyGluPheGlnHisLeuGlnCysSerAlaArgValSerIleAsnThrValAsn 200
 Db 605 GGTCAATTTCCNGAATTTGGGACGATTTGTCAGTGTCTGACAGATCTCTGACACCAACAT 664
 QY 201 LeuTrpValCysProGlnValMetGluValLeuPheArgTrpHisGlyArgAlaTyr 720
 Db 665 GTCACTCTGGGCTCACTCACTGAAGATGACTGTCTACAAACACATGGACGGCGATAT 724
 QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
 Db 725 GTVCCATC-GSACAGCTGGAAAT-GTATACCTGGTGTMA-GATCARAT-CTGTGTGTGTG 780
 QY 241 ThrMetTyrGlnLysAspArgAsnSerSerAspGlyThrPheLeuArgAspLeuPro 260
 Db 781 ---CTATGTCACGAAATATGAAATCATTCGCGGAACCTCTCAAGATCT---CCC 834
 QY 261 IlePhePheAspVal 265
 Db 835 ATATGTTTGATGTC 849

RESULT 13
 BB613060 LOCUS
 DEFINITION 623 bp mRNA linear EST 31-AUG-2001
 musculus cDNA clone 4732452017 5', mRNA sequence.
 BB613060 1 GI:15395417
 SOURCE house mouse
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 623)
 Authors Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 O., Shibata, K., Shimizu, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki
 D., Shibata, K., Shinozaki, T., Saito, R., Saito, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, P., Tanaka, F., Tanaka, T., Toyota, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN
 Genomic Science Institute of Physical and Chemical
 Institute of Physical and Chemical Research (RIKEN)
 1-7-29, Suwayama, Tsukuba, Ibaraki, 305-8565, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10): 1617-1630 (2000)
 Watanabe, K., Fujita, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanabe, M., Yoneda, Y., Ishikawa, I., Ozawa, K., Tanaka, T., Matsura

Search completed: April 9, 2003, 08:46:38
Job time : 2488 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 8, 2003, 10:06:23 ; Search time 4818 Seconds

(Without alignments)

3455,128 Million cell updates/sec

Title: US-09-943-075a-2

Perfect score: 30.01

Sequence: 1 MESLCCVLLVLLAAGLPLQ.....PFSRGREKDPILLQKPNKL 572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xsepxt 0.5
Ygapop 10.0 , Ysepxt 0.5
Fgapop 6.0 , Fsepxt 7.0
Delop 6.0 , Delxpt 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=wellh
-O=/cnp2.1/USPTO-spec01/US05943075/runat_07042003_155630_17927/app_query_faeta.1.711
-DB=GenEmbl -OFMT=FASTap -SUPERIX=Ige -MINMATCH=0.1 -LQOPCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40_cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US05943075.ecom_l_15586-trinet_07042003_155630_17927 -NCPU=6 -ICPU=3
-N=1 -NTHREADS=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FONOP=6 -FONEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb.ba:.*
2: gb.htg:.*
3: gb.in:.*
4: gb.in:.*
5: gb.ov:.*
6: gb.pat:.*
7: gb.ph:.*
8: gb.pl:.*
9: gb.pr:.*
10: gb.ro:.*
11: gb.ro:.*
12: gb.sv:.*
13: gb.un:.*
14: gb.vl:.*
15: em.ba:.*
16: em.fun:.*
17: em.hum:.*
18: em.in:.*
19: em.in:.*
20: em.em:.*
21: em.or:.*
22: em.ov:.*
23: em.pat:.*
24: em.ph:.*
25: em.pl:.*
26: em.ro:.*
27: em.ro:.*
28: em.un:.*

29: em.vl:.*
30: em.htg_hum:.*
31: em.htg_inv:.*
32: em.htg_other:.*
33: em.htg_mus:.*
34: em.htg_pin:.*
35: em.htg_rna:.*
36: em.htg_sap:.*
37: em.htg_vrt:.*
38: em.sy:.*
39: em.htgo_hum:.*
40: em.htgo_mus:.*
41: em.htgo_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3061	100.0	2320	10	AF184983 Rattus no
2	2705	88.4	2305	10	MMU251685
3	2703	88.2	2213	6	ARI56839 Sequence
4	2703	88.2	2279	10	AF322054 Mus muscu
5	2163	70.7	2282	10	AF322055 Mus muscu
6	2163	70.7	2282	6	AF322056
7	2163	70.7	2669	6	AR018808 Sequence
8	2163	70.7	2669	6	AR018808 Sequence
9	2163	70.7	2669	9	HSNMB
10	2162	70.6	2683	6	AX358788 Sequence
11	2162	70.6	2683	6	AX358788 Sequence
12	1461.5	46.4	2467	5	C3ON77
13	1461.5	46.4	2467	5	C3ON77
14	941	30.7	620	6	AR018825
15	870	28.4	162541	2	AC098114
16	823	26.9	104229	2	AC128735
17	816	26.7	190519	2	AC110271
18	710.2	23.8	1524	6	AR018817
19	710.2	23.8	1524	6	AR018817
20	711	23.4	1473	6	AR018827
21	640.5	20.9	2470	5	DB8348
22	629	20.5	2026	9	HS001874
23	627.5	20.5	2114	9	HUMGPMSS
24	627.5	20.5	2115	6	A45993
25	627.5	20.5	2130	6	AR167365
26	627.5	20.5	2130	6	AR167365
27	627.5	20.5	2130	6	AX354933
28	627.5	20.5	2130	9	S73003
29	627.5	20.5	2134	9	BC001414
30	627.5	20.5	2134	9	BC001414
31	627	20.5	2172	6	AR063067
32	627	20.5	2172	6	AR063067
33	627	20.5	2172	6	AR063067
34	627	20.5	1992	6	AR162397
35	622.5	20.3	1986	6	AX133655
36	622.5	20.3	1986	6	AX133655
37	622.5	20.3	1986	6	AX192347
38	622.5	20.3	2534	6	AX133528
39	622	20.3	2131	6	AX474662
40	620	20.3	1881	6	HUMPMEL
41	6120	20.3	1881	10	MMU14133
42	611.5	20.0	18979	2	AC003082
43	611.5	20.0	18979	2	AC003082
44	524	17.1	335	6	AR018809
45	524	17.1	335	6	AR018811

ALIGNMENTS

RESULT 1

Dn	751	ATTCCTCATCTCGAAGGTGAAAGTGTGTATGTGATACAGACAGATCCCTGTATTCTGT	810
Qy	241	ThrMetTyrGlnIysAsnAspAsnSerSerAspGluThrPheAsnAspLeuPro	260
Dn	811	ACCATGTGCCAGAGAAATGACAGAACTTCTGTATGATGATCTTCTCAGACACTCCCC	870
Qy	261	IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer	280
Dn	871	ATGCTGTTCGATGCTCATATGATGATCCAGCAACTCTCCTCAAGCACTTCCATTTCC	930
Qy	281	TyrLysTyrAspPheGlyAspAsnThrGlyLeuPheValSerAspHisIleThrLeu	300
Dn	931	TACAAGTGGAACTTTGGGACACACCTGGCTGTTGTCTCCACAAATCAGACTTTGAAT	990
Qy	301	HisThrTyrValLeuAsnGlyThrPheAspPheAsnLeuThrValGlnThrAlaValPro	320
Dn	991	CACACTTATGTCTCATGAACTTTCACCTTAACTTACCCTGACAGCTGACAGTCCCC	1050
Qy	321	GlyProCysProSerProThrProSerProSerSerSerSerProSerPro	338
Dn	1051	GGGCGATCGCTCCCTCTCGCTTCTGCATCGCGCTTCAACCTCCAGCTCCGCTTAACT	1110
Qy	339	AlaSerProSerProThrLeuSerThrProSerProSerSerLeuMetProThrGlyTyr	358
Dn	1111	TGGCTCTACCTTTCGCACATATACACCTTACCTTCTTAAATGCTACCTGTATC	1170
Qy	359	LysSerMetGluLeuSerPheIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr	378
Dn	1171	AAATCCATGGAGCTGAGTGACATTTCCAAATGAACTCCCGAATACAGATATGGCTAC	1230
Qy	379	PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleValAla	398
Dn	1231	TTCAAGCCACATCAATATGAGAGGATCTCGAAGTCAAGTCACTACGATGATGCA	1290
Qy	399	AspValProIleProThrLeuGlnProAspAsnSerIleuMetAspPheIleValThrCys	418
Dn	1291	GATCTCCCATGCCACACCGACCTCCCAACTCTCGATGGACTTCACTGTGACCTGC	1350
Qy	419	LysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAla	438
Dn	1351	AAAGGGGCGACCCCATGGAAGCTGTACGATCATCTCCGACCCCACTGGCAGTCCG	1410
Qy	439	AlaAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArg	458
Dn	1411	CAGAACGGGTGTCGACCTCTGGCTGGTGGATGGGCTGCTGCTGTCTGTGAGAGA	1470
Qy	459	AlaPheAsnGlySerGlyThrCysValAsnPheThrLeuGlyAspAspAlaSerLeu	478
Dn	1471	GCTTTCATGGGTCTGGCACTACTGTGAATTTCACTCTGGAGATGATGCAAGCGTG	1530
Qy	479	AlaLeuThrSerAlaLeuIleSerIleProGlyIysAspLeuGlySerProLeuArgThr	498
Dn	1531	GCTCTCACACACCTGATCTATCTTCCTTGGCAAGCCACCACTCCCTCTGAGAGA	1590
Qy	499	ValAspGlyLeuLeuIleSerIleGlyCysAlaMetPheValThrMetValThrIle	518
Dn	1591	GTGAATGTGTCTCACTCCATCGCTGCTGCTGTCTGTCTGTCACTCATCTGTTACATC	1650
Qy	519	LeuLeuTyrIysIysHisIysThrTyrIysProIleIleIysAsnCysThrArgAsnVal	538
Dn	1651	TTGCTGTCAACAAACACAGGCGTACAGCCATAGAAACTTCCCCAGACAGCGTCT	1710
Qy	539	LysGlyIysGlyLeuSerValPheLeuSerHisAlaValAlaProPheSerArgGlyAsp	558
Dn	1711	MAAGGCAAGGCGCTGAGTGTCTCTCCTCAGTCAGCGAAGGCCCGCTTCTTCGAGGAG	1770
Qy	559	ArgGluIysAspProLeuLeuGlnAspIysPro	569
Dn	1771	CAGGAGAGGATCCATTTGCTCCAGGACAGCCA	1803
LOCUS	ARI56839	2213 bp	DNA linear PAT AUG-2001

DEFINITION	Sequence 27 from patent US 6242419.
ACCESSION	ARI56839
VERSION	ARI56839.1
KEYWORDS	GI:15125543
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	Strachan,L., Sleeman,M., Abernethy,N., Onrust,R., Kumble,A. and Arslan, C.
TITLE	Compositions Isolated from stromal cells and methods for their use
JOURNAL	Patent: US 6242419-A 27 05-JUN-2001.
FEATURES	Location/Qualifiers 1..2213 1..2213 /organism="unknown"
BASE COUNT	526 a 559 c 533 t
ORIGIN	
Alignment Scores:	
Pred. No.:	6.2e-175 Length: 2213
Score:	2700.00 Matches: 507
Percent Similarity:	92.29% Conservative: 20
Best Local Similarity:	88.79% Mismatches: 42
Query Match:	86.21% Indels: 2
Dn:	Gaps: 1
US-09-943-075a-2 (1-572) x ARI56839 (1-2213)	
Qy	1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
Dn	23 ATGGAACTCTCGGGCTCTGGGATTTTCGCTGGCTGCAGGATGCTCTTCAG 82
Qy	21 AlaAlaIysArgPheAspAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Dn	83 GCTGCCAAGGATTTCTGATGTGCTGGCCATCAACAGTATCCCAATCAGATGAGAG 142
Qy	41 AsnAsnGlnLeuArgGlyTyrPheSerAspGluAsnGluThrPaspGluLeuTyrPro 60
Dn	143 CACACCACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 200
Qy	61 ValTyrPArgGlyGlyGlyValGlyTyrPheAspSerTTPGlyGlyValGlyValGlnAla 80
Dn	203 GTGTGGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 262
Qy	81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Dn	263 CTCTGTCACAGTACCTACCGGCTCTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
Qy	101 ValPheProArgCysGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyG 120
Dn	323 GTGTGTCCTCCAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 382
Qy	121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnThrThrThrThrThr 140
Dn	383 AGAAATGATTTGGGACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
Qy	141 AspGluAspArgPheGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGly 160
Dn	443 GATGTGTGATCTGGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502
Qy	161 ProPheProArgProHisGlyArgIysTyrAsnPheValTyrValPheIleThrLeu 180
Dn	503 CCCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 562
Qy	181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Dn	563 GGCCAGTATTTCCAAAACCTGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 622
Qy	201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220
Dn	623 TTGACAGCTGGCTGCTGAGGTCATGGAGTGCATCTCTTTCGAGATACGGCTGGGCA 682
Qy	221 IleProIleSerLysLysAspValTyrValIleThrAspGlnIleProIlePheVal 240

[illegible]

Db 1130 CTGAGCTGAGTACGATTCCTGAGGAACTGCCAGATATACAGATATGCGCACTTCAC 1189
 Qy 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
 Db 1190 GCCACCATCAATATTGAGAGGAATCTTAGAGTCTTACATCATCAACGACAGACGTC 1249
 Qy 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
 Db 1250 CTGATGCGTGGCCGATCGGCTACAGAACCCCTTAATAGACTTGTGTGACCTGCCAAGG 1309
 Qy 421 AlaThrProGlnIleValThrIleLeuSerAspProThrCysGlnIleAlaGlnAsn 440
 Db 1310 ASCATTCCACGAGGCTGTACCATCATTTCTGACCCCACTCGCGAGATCACCAGAAC 1369
 Qy 441 ArgValCysSerProValAlaValAspIleLeuCysLeuSerValArgArgAlaPhe 460
 Db 1370 ACAGCTCGACGCCCTGTGATGCTGCAGAGACCCAGCCAGCTGCGCTTAGAGTGGCAAC 1429
 Qy 461 AsnGlySerGlyThrIleCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
 Db 1430 AATGGTCTGGAGCTACTGTGTGAACCTCACCTGGGGGATGACAAAGCTGGCTCTC 1489
 Qy 481 ThrSerAlaLeuIleSerIleProGlyIleAspLeuGlySerProLeuArgThrValAsn 500
 Db 1490 ACAGACCCCTGATTTCTCTGCAGAGACCCAGCCAGCTGCGCTTAGAGTGGCAAC 1549
 Qy 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValMetValThrIleLeuLeu 520
 Db 1550 AGTCGCTGATCTCCGCTGGCTGTGGCTATGCTGATTTGTCACTGTGATCCCTTTGGT 1609
 Qy 521 TyrIleValHisLysThrIleProIleGlyAsnCysThrArgAsnValValLysGly 540
 Db 1610 TCAAGACACACAGAAATACACCAATAGAAATAGTCTGGGATGTGTGACAGAC 1669
 Qy 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgLysAspArgLys 560
 Db 1670 AAGGCGCTGAGTGTCTTCTCAACGCTGCMAAAGCCGTGTCTCCCGGAAACACGAA 1729
 Qy 561 LysAspProLeuGlnAspLys 568
 Db 1730 AAGATCCGCTACTCAAAACCA 1753

RESULT 8
 LOCUS AR018814 2669 bp DNA linear PAT 05-DEC-1998
 DEFINITION Sequence 101 from patent US 578182.
 FEATURES
 CDS
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE
 1. (bases 1 to 2669)
 Thompson, T. C.
 Method for identifying metaabiotic sequences
 Patent: US 578182, 1998, JUL-1998;
 TITLE Location/Qualifiers
 source 1..2669 /organism="unknown"
 BASE COUNT 752 a 595 c 597 g 725 t
 ORIGIN

Alignment Scores:
 Pred No.: 3.14e-138 Length: 2669
 Score: 2163.00 Matches: 394
 Percent Similarity: 82.57% Conservative: 75
 Best Local Similarity: 69.37% Mismatches: 85
 Query Match: 14
 DB: 6 Gaps: 1

US-09-943-075a-2 (1-572) x AR018814 (1-2669)
 Qy 1 MetGluSerCysGlyValLeuValPheMetLeuAlaAlaGlyLeuProLeuGln 20

Db 92 ARGAGATGCTCTCACTATTCTCCGGGATTTCTCTGGCTGCGAGATGCCACTTGAT 151
 Qy 21 AlaAlaLysAspArgPheArgAspValLeuGlyHisGlnArgGlu 211
 Db 152 CGCGCCAAACGATTTTCATGTATGCTCGGCCAATCAAGACACTTCTGCTTACATGAGGAG 400
 Qy 41 AsnAsnGlnLeuArgGlyIlePheSerSerAspGlnAsnGlnIlePheValValLeuIlePro 60
 Db 212 CACATCATTAATAGCTGTCTCTGATGAAATGACTGGAATGAAATGAAATGCTATCCCA 271
 Qy 61 ValITrArgArgGlyGlnGlyArgGlyIlePheSerSerPheGlyGlyGlyValGlnAla 80
 Db 272 GTGTGAAGCGGGAGACATGAGGTGGAAACCTCTCGAAGSGAGCGCTGTGCGAGCG 331
 Qy 81 AlaLeuThrSerAspSerProAlaLeuValValGlySerAsnIleThrPheValValAsnLeu 100
 Db 323 GTCTCGACAGACGACTCACAGAGCCCTGTGGGCTCAAAATACATTTTCGGTGAACCTG 190
 Qy 101 GlyProCysArgGlyGlnGlyAsnThrSerGlnGlyHisLeuValPheValPheAsnGly 320
 Db 392 ATATTCCCTAGATGCCAAGAGAGATGCCAATGTCACATAGTCTATGAGAAGAACTGC 451
 Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnThrThrGlyAlaAsp 140
 Db 452 AGAATAGAGCTGGTATTATCTGCTGCTCCATGCTGTTCACACTGACATGTCACACT 511
 Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlyHisLeuValPhePheAspGlyLys 160
 Db 512 CACATGACGCGGGAATAATGCCCGCCAAAGCCATCAATAAGCTCTTCCTGATGGGAA 571
 Qy 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180
 Db 572 CTTTCTCTACCCACCCGCTGGAGARGTGGATTTCACTGCTGCTCCACACT 631
 Qy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
 Db 632 GGTCAAGTATTTCCAGAAATTTGGGACGATGTTCAAGTAGAGTTCTGTGACACACCAAT 691
 Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220
 Db 692 GTGACACTTGGGCTCTCATCTGGAAGTGCATCTACAGAGAGAGATGAGACGGCATAT 751
 Qy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
 Db 752 GTTCCCATCCGCAACATGAAGATGCTGACGTGTGACGATCAGATCAGATCTGTGTGTG 811
 Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
 Db 812 ACTATGTTCCAGAGAGATGAAATTCATCCGACGAACCTCTCTCAAGATCTCC 280
 Qy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
 Db 872 ATATTGTTGATGCTCATGATGATGCTGACCTAGCCACTTCTCAATTTCTTACCATTAAC 931
 Qy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
 Db 932 TACAGTGGAGCTTCGGGATTAATCTGCGCTGTGTGTTCCCACTACTACTGTGAT 991
 Qy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnIleThrAlaIlePro 320
 Db 992 CACAGTATGTGCTCAATGGAACTCTGACCTTAACCTCACTGTGAAAGCTGACGACCA 1051
 Qy 321 GlyProCysProThrProSerProSerSerSerSerSerSerSerSerSerSerSerSer 340
 Db 1052 GGACCTTTCGCCACCCACCCACCCACCCACCCACCT 360
 Qy 341 SerProSerProThrLeuSerProSerProSerProSerProSerProSerProSerPro 380
 Db 1088 -----TCAAAACCCACCCCTCTTTAGACCTGCTGGTGAACACCC 1129
 Qy 361 MetGluSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380

Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220
 Db 692 GTACACTTGGGCTTCACTCATGAGTACTGCTACAGAGACAGGCGGCTAT 751
 Qy 221 IleProIleSerValLysAspValTyrValIleThrAspGluIleProIlePheVal 240
 Db 752 GTTCCATCCGCAAGTCAAGATGTGTAGCTGGTAAACAGTCAAGATTCCTGTGTGTG 811
 Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
 Db 812 ACTATGTTCAGAGAGAGTCAAGATTCATTCGAGAGACCTTCTCAAGATCTCCCC 871
 Qy 261 IlePhePheValIleIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
 Db 872 ATATGTGTGATGCTGATTCATGATCATGATCTAGCAGCTCTCTCAATATTCACCATTAAC 931
 Qy 281 TrrTyrTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisIleThrLeuAsn 300
 Db 932 TACAGTGGACCTTCGGGGATATTCGCTGTGTTGTTCCACCAATCACTACTGATGCT 951
 Qy 301 HisThrTyrValLeuLeuIleThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 Db 952 CACAGTATGTCTTAATGATGACCTTCAGCTTACCTCACTGTAAGCTGGACAGCA 1051
 Qy 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340
 Db 1052 GGACCTTGTCCGCGACCGCCACCCACCCAGACT----- 1087
 Qy 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrLysSer 360
 Db 1088 -----TCAAAACCCACCCCTCTTATAGGACCTGCTGTGCACACCCC 1129
 Qy 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380
 Db 1130 CTGGAGCTGAGTAGGATCTGATGAAACCTCCAGATTAAACAGATATGCGCCACTTCAA 1189
 Qy 381 AlaThrIleThrIleValAspGlyIleLeuIleValAsnIleIleIleIleIleIleIleIleIle 400
 Db 1190 GGCACCTCATATGTAGAGGAACTTGTAGTGTAACTCATCTCAGATGACAGATC 1249
 Qy 401 ProIleProThrLeuGlnProAsnSerLeuMetAspPheIleValThrCysArgGly 420
 Db 1250 CTGATCGCGGCGCCATGGCTGAAGCTCCCAATAGACTTGTGTGCTGACCTGCCAAGG 1309
 Qy 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysIleIleIleIleIleIle 440
 Db 1310 AGCATCCGACAGGCTGTACACTTCTGATGCCCTCTGGAGATCAACCAAGAC 1369
 Qy 441 ArgValCysSerProValAlaValAspGluLeuLeuLeuSerValArgAlaPhe 460
 Db 1370 ACAGTCTGCAGCCCTGTGATGTGGATGATGTGTCTGCTGATGTGAGAGAACCTTC 1429
 Qy 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
 Db 1430 AATGGCTTGGGAGTACTGTGTGACCTCACCTGGGGATGACACACCTGGCTCTC 1489
 Qy 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
 Db 1490 AGCAGACCCCTGATTTCTGCTGCAGAGACCCAGCCCTGCTTTAAGGATGGCAAC 1549
 Qy 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520
 Db 1550 ATGTGCCCTGATCTCGCTGGCTGTGGCAATTTGTCATCTGATCTCCCTCTGGTG 1609
 Qy 521 TrrTyrLysIleHisLysIleThrTyrLysProIleCysIleAsnCysThrArgAsnValValLysGly 540
 Db 1610 TACAAAAACACAGGAAATACACCCATAGAAAATAGTCTGGGAATGTGTGTCAGAG 1669
 Qy 541 LysGlyLeuSerValIlePheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
 Db 1670 AAGGCGTGTGTGTCTTCTCATCCCTGCAGAAAGCCCTGCTTCTCCGGAGAACAGAGA 1729

Qy 561 LysAspProLeuLeuGluAspLys 568
 Db 1730 AAGGATCGGCTACTCANAAACCA 1753
 RESULT 10
 LOCUS AX358788 2683 bp DNA linear PAT 13-FEB-2002
 DEFINITION Sequence 41 from Patent WO0193983.
 ACCESSION AX358788
 VERSION AX358788.1 GI:18675295
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 AUTHORS Baker,K.P., Desnoyers,L., Gerritsen,M.E., Goddard,A.,
 Godowski,P.J., Grimaldi,J.C., Gurney,A.L., Smith,V., Stephan,J.P.,
 Szeles,C.C., and Weinreb, R.
 TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
 the same
 JOURNAL Patent: WO 0193983-A 41 13-DEC-2001;
 Genentech Inc. (US)
 FEATURES
 source 1. 2683 bp
 /db_xref="taxon:9606"
 BASE COUNT 751 a 504 c 596 g 732 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,76-138 Length: 2683
 Scores: 216.00 Acceses: 76
 Percent: 82.72% Consistative: 78
 Best Local Similarity: 69.1% Mismatches: 87
 Query Match: 70.63% Indels: 12
 DB: 6 Gaps: 2
 US-09-943-075a-2 (1-572) x AX358788 (1-2683)
 Qy 1 MetGluSerLeuCysGlyValIleValPheLeuLeuAlaGlyLeuProLeuGln 20
 Db 82 ATGCAATGCTCTACTATTTCTCGGGATTTCTCTCTGCTGCAAGATGCCACTTGCAT 141
 Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
 Db 142 GCCCCCAACATTCATGTGTGGCAATGACAGACCTTCGCTTCATCAGAGG 201
 Qy 41 AsnAsnGlnLeuArgGlyTyrPDSerSerAspGluAsnGluTyrAspGluLeuTyrPro 60
 Db 202 CACATCAATTAATAGCTGTGCTTCTGATGAAATGACTGGAATGAATACTACCCA 261
 Qy 61 ValTrpArgArgGlyGluGlyArgTyrTrpLysAspSerTrpGluGlyArgValGlnAla 80
 Db 262 GTGTGAGGAGGGGAGACATGTGTGTGGAAATCTCTGTAGAGAGCCGTGTGCAGCC 321
 Qy 81 AlaLeuThrSerProAlaLeuAlaGlySerAsnIleThrPheValValAsnLeu 100
 Db 322 GTCTGCACCATGACTCACCGACCTCTGGGCTCAATATACATATTTCGGGTGAACCTG 381
 Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
 Db 382 ATATTCCTAGATGCCAAGAGAGATGCCATGGCAATAGCTATGAGAGACCTGC 441
 Qy 121 ArgSerAspLeuGluAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140
 Db 442 AGAATAGAGCTGGTGTATCTGCTGATCGCTATGTTTACACTTGCACACATGGTCAGAG 501
 Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
 Db 502 GACATGTACCGGGAAGATGGCCCGCCGCAACCCATCAACCTCTTCCCTGATGGAGA 561
 Qy 161 PropheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180

```

Db 562 CTTTTCCTCCACACCCCGATGGAGAGATGAAGTTCATCTACGCTCTCCACACACTT 621
Qy 181 GUGUATrVpHeGluLeuGluGlyCysSerAlaArgValSerIleAsnThrPro 200
Db 622 GGTCAATGATTCACGAAATTCGGAGCATCTTCAGCTGAGATTTCTGTGAACACACCAAT 681
Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgAArgHisGlyValGlyAlaTyr 220
Db 682 GTGACACTTGGGCGCTCACTCATGATGAAGTGAAGTCTCTACAGAACATGAGCGGGCATAT 741
Qy 221 IleProIleSerValValLeuValAspValTyrValIlePheAspGlnIleProIlePheVal 240
Db 742 GTTCCCATCGCAGCAAGTCAAAAGTGTGTACCTGGTAAACATCAGATTCCTGTGTGTGTG 801
Qy 241 ThrMetCtYrGlnLysAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
Db 802 ACTATGTCTCGAAGACAGATCGAAATTCATCCGCGAAACCTTCTCGAAGACCTCCCC 861
Qy 261 IlePhePheAspValLeuIleLeuAspProSerGlnPheLeuThrValSerAlaIleSer 280
Db 862 ATTATGTTGTATGCTCTCATTCATGATCTAGCTTACGCTTCTCCATATATGCTACCATAC 921
Qy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
Db 922 TACAAAGTGAGCTTCGGGATTAATCTGGCGCTGTGTTTCCACCAATCATACTGTGAAT 981
Qy 301 HisThrTrpValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Db 982 CACAGTATGTCTCACTGCACTGCAACCTTCACTGCTTACCTGACTGTGAAGCTGCAGCACCA 1041
Qy 321 GlyProCysProSerProThrProSerProSerSerThrSerProSerProAlaSer 340
Db 1042 GACACTGTCTCGCCACCCACCCACACCCACAGA-----CCTTCA 1080
Qy 341 SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro 355
Db 1081 AAACCCACCCCTCTTTAGCACTACTCTAAATCTTATGATTCAAACACCCCGAGCCT 1140
Qy 356 ThrGlyTyrIleSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg 375
Db 1141 CAGTGTACAGAGCTCTGATGCGGTGCTGCGCTGGAAGCTCTCTAATAGACTTTGC 1320
Qy 416 ValArgAlaValPheAsnGlySerGlyThrTrpCysValAsnPheThrLeuGlyAspAsp 435
Db 1321 GTGAGCTGCGCAAGGAGCATTCACACGAGGCTGTGTACCATCTTCTGACCCCACTGCG 1380
Qy 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer 455
Db 1381 GAGTACCAAGACAGTCTGACGCTCTGGATGGGTGAGATGGTGTGCTGACT 1440
Qy 456 ValArgAlaValPheAsnGlySerGlyThrTrpCysValAsnPheThrLeuGlyAspAsp 475
Db 1441 GTGAGAGCAACTCTCAATGGGTGTGGAGCTACTGTGTGAACCTCACTCGCGGGATGAC 1500
Qy 476 AlasrLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerPro 495
Db 1501 ACAAGCTGGCTGTACAGAGACCCCTGATGTGTTCTGACAGAGCCACCCCTGCTCT 1560
Qy 496 LeuAlaGthrValIleAlaCysValLeuIleSerIleCysValAlaMetPheValThrMet 515
Db 1561 TTAAGGATGGCAACAGTGGCCCTGATCTCCGCTGGCTGCTGGCCATATTTGTCACTGTG 1620
Qy 516 ValThrIleLeuLeuTyrLysLysHisLysThrTyrTyrProIleGlyAsnCysThrArg 535

```

```

Db 1621 ATCTCCCTCTTGGTGTACAAAAACAGAGATCAACCCCATAGNAATATGCTCGG 1680
Qy 536 AsnValValIysGlyGlySerValPheLeuSerHisAlaLysAlaProPheSer 555
Db 1681 AATGTGTGAGACAGAACGCTGTGCTCTTCACAGCTGGACAGAGCGTGTCTTTC 1740
Qy 556 ArgGlyAspArgGlyLysAspProLeuLeuGlnAspLys 568
Db 1741 CGSGMAACACGAGAAAGGATCCGCTACTCAAAACCAA 1779

RESULT 11
ACCESSION AK362281
DEFINITION Sequence 41 from Patent WO0208288.
ACCESSION AK362281
VERSION AK362281.1 GI:18694598
KEYWORDS human.
SOURCE human.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
  Baker,K.P., Desnoyers,L., Gerritsen,M.E., Goddard,A.,
  Godowski,P.J., Grimaldi,J.C., Gurney,A.L., Smith,V., Stephan,J.P.,
  Watanabe,C.K. and Wood,W.I.
  Secreted and transmembrane polypeptides and nucleic acids encoding
  the same
  Patent: WO 0208288-A 41 31-JAN-2002;
  Genbank: F0208288.1
FEATURES
  Location/Qualifiers
    source
      1..2683
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT 751 a 604 c 596 g 732 t
ORIGIN

Alignment Scores:
  Pred. No.: 37e-138 Length: 2683
  Score: 2162.00 Matches: 396
  Percent Similarity: 82.72% Conservative: 78
  Best Local Similarity: 69.11% Mismatches: 87
  Query Match: 70.63% Indels: 12
  Db: 6 Gaps: 2

US-09-943-075A-2 (1-572) x AK362281 (1-2683)
Qy 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20
Db 82 ATGGATGCTCTACTATTCTCTGGGATTTCTCTCGCTGGCTGAAGATGCACTGTAT 141
Qy 21 AlaAlaValGlyPheArgValValLeuGlyHisGluGlyTyrProAspHisMetArgGlu 40
Db 142 GCGCCCAACGATTTTCATGATCTGTCTGGCAATGAAGACCTTCTCTTACATGAGGAG 201
Qy 41 AsnAsnGlnLeuArgGlyTyrPheSerAspGluAsnGluTyrPheLeuLeuTyrPro 60
Db 202 CACATCAATTAATGCTGGTCTGTCTGTATCAAAATGACTGGAATGAACACTTACCCA 261
Qy 61 ValTrpArgGlyGluGlyValPheAspSerTrpGluGlyValGlyValGluAla 80
Db 262 GTGTGGACCGGGGAGACATGAGGTGGAAAAAACTCTGGAAGCGGCGCTGTGCGGCG 321
Qy 81 AlaLeuSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 322 GTCTTGACCACTGACTCACACGCGCTCTGGGGCTCAATATATACATTTTGGGGTGA 381
Qy 101 ValPheProArgCysGlnIleValGlyAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
Db 382 ATATTCTCTAGATGCGAAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGACTGC 441
Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnIleThrThrGlyAlaAsp 140
Db 442 AGAATGAGCTGGTGTATCTCTGCTGCTGATGTTTACACACTGGACGATGCTCAGAG 501

```

QY 141 AspGluAspTTPgluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLeu 160
 DB 502 GACAGTCACGGGGAAATGCGCCAAAGCCCAACATCACTACGCTCTCCCTGATGGAAA 561
 QY 161 ProPheProArgProHisHisGlyArgLysLysIleValPheHisThrLeu 180
 DB 562 CTTTTCCTCCACACCCCGGATGAGAGATGAATTCATCTACGCTCTCCACACACT 621
 QY 181 GlyValIleThrPheGlnLysLeuGlnCysSerAlaArgValSerIleAsnThrValAsn 200
 DB 622 GGTCAAGTATTTCCAGAAATGSGACGATGTCAGTAGAGATTTCTGTGACACAGCAAT 681
 QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220
 DB 682 GTGACATGTGGGCTCACTCATAGAGATGACTGTCTACAGAGATCGGGGCGCAT 741
 QY 221 IleProIleSerLysValLysValAspValTyrValIleThrAspGlnIleProIlePheVal 240
 DB 742 GTTCCCATCGCAAGTGAAGATGTGTAGCTGTGTACACAGATCAGATTCCTGTGTGTG 801
 QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspIleThrPheLeuArgAspLeuPro 260
 DB 802 ACTATGTTCGAGAGACATGATGATCATCGAGAGACCTCTCTCAAGATCTCCCG 861
 QY 261 IlePhePheAspValIleuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
 DB 862 ATATGTTTGAATGCTGATCATGATCATGATCGTGTGTACACAGATCAGATTCCTGTGTGTG 921
 QY 281 TyrIleGlnPheGlnAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300
 DB 922 TACAAGTGGAGCTTCGGGATATATCTGGCTGTGTGTGTTCACCATCATCATCTGTGAT 981
 QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValClnThrAlaValPro 320
 DB 982 CACAGCTATGTGCTCAATGAACTCTACAGCTCTACCTCAGCTGTGAAGCTGCACACCA 1041
 QY 321 GlyProCysProSerProThrProSerProSerSerThrSerProSerProAlaSer 340
 DB 1042 GGACCTGTGTGGCCACCCACCCACCCAGAC-----CCTTCA 1080
 QY 341 SerProSerProThrLeuSerThr-----ProSerProSerProSerProSerPro 355
 DB 1081 AAACCCACCCCTCTCTTATGACCACTCTAAATCTTATGATTCACACCCACCGAGACT 375
 QY 356 ThrGlyTyrLysSerMetGluLeuSerSerHisSerAsnGluAsnCysArgIleAsnArg 395
 DB 1141 ACTGTGTGAACCCCTGGAGCTGAGTAGGATCTCTGTATGAATGCCAGATTAACAG 1200
 QY 376 TyrGlyTyrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle 395
 DB 1201 TATGGCCCTCTTCAAGCCACACATCAATGTGTAGAGGGATCTTAGAGGTATACATCATC 1260
 QY 396 GlnValAlaValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
 DB 1261 CAGATGACAGAGCTCTGTATGCGCGGTGCGCATCGCTGAAAGCTCTCAATAGATTCTTC 1320
 QY 416 ValThrCysLysIleAlaThrProThrGluAlaCysThrIleSerAspProThrCys 435
 DB 1321 GFGACCTGCAGAGGAGCATCTCCAGAGAGCTGTGTACCATCATTTCTGACCCACCTCG 1380
 QY 436 GlnIleAlaGlnAsnArgValCysSerProGlnAlaValAsnGluLeuGlyLeuSer 455
 DB 1381 GAGATCAACCCACAGACAGCTCTGACGCGCTGTGGAGTGGATGATGTCTGTCTGACT 1440
 QY 456 ValArgAlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAsp 475
 DB 1441 GFGACAGACACTCTCAATGGCTGTGGAGCTACTGTGTGAACCTCACTCCCTGGGGATGAC 1500
 QY 476 AlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerPro 495
 DB 1501 ACAAAGCTGGCTCTCACAGACACCTGATTTCTGTCTGACAGACAGACCCAGCTCGCT 1560

QY 496 LeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMet 515
 DB 1561 TTAGAGATGCGAAGAGTCCCTGATCTTCGTGGCTGCTGTGGCATTTGTTCATCTGTG 1620
 QY 516 ValThrIleLeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArg 535
 DB 1621 ATCTCCCTCTCTGTGTACAAAAACACAGGAATACACCACTAGAAAATAGTCCCTGGG 1680
 QY 536 AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSer 555
 DB 1681 AATGTGTGTGAGAGAGGCTGTAGTGTCTTCACCTGTGCAAAAGCCGTGTCTTC 1740
 QY 556 ArgGlyAspArgGluLysAspProLeuLeuGlnAspLys 568
 DB 1741 CCGGAAACACGAGGAAGATCGCTACTCANAAACCAA 1779
 RESULT 12
 LOCUS BC032783
 DEFINITION Homo sapiens, similar to glycoprotein (transmembrane) nmb, clone
 BC032783
 VERSION BC032783.1 GI:21619752
 FEATURES
 CDS
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Genome Sequencing Center, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 DNA Library Preparation: Life Technologies, Inc.
 Cloning Vector: pCMVSPORT6
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland.
 Web site: <http://www.nih.gov/>
 Contact: nisc_mgc@nih.gov
 Author: N. Ayele, K. Beckstrom-Sternberg, S. M., Benjamin, B.,
 Dierich, N. L., Granite, S. Guan, X., Gupta, R., Hsieh, C.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q. L., Masilelo, C., Maskeri, B., Mastrian, S. D., McCloskey, J. C.,
 McDowell, J., Pearson, R., Santipol, S., Thomas, P. J., Touchman, J. W.,
 Tsaurgon, C., Vost, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L.,
 Young, A., Zhang, L. H., and Green, E. D.
 Clone distribution: MGC clone distribution information can be found
 through the T.M.A.G.E. Consortium/GENI at: <http://image.llnl.gov>
 Series: IRAC Plate: 68 Row: k Column: 7
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4505404.
 Location/Qualifiers
 1. 2787
 /organism="homo sapiens"
 /db="genbank"
 /clone="MGC:45334 IMAGE:5177095"
 /tissue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 /lab_1796
 /lab_1797
 /product="similar to glycoprotein (transmembrane) nmb"
 /protein_id="AAH32783.1"
 CDS
 FEATURES
 source

FEATURES		Location/Qualifiers
1. .620		
BASE COUNT 143 a 207 c 128 g 142 t		
ORIGIN		1. .620
Alignment Scores:		
Pred. No.:		9.57e-56
Score:		941.00
Serec. Similarity:		91.75%
Base Similarity:		91.75%
Query Match:		30.74%
DB:		6
Length:		620
Matches:		183
Conservative:		6
Mismatches:		15
Indels:		3
Gaps:		1
US-09-943-075a-2 (1-572) x AR018825 (1-620)		
Db	942	GGAATGGCTTATTGTGACAAAGCGCGCTACTACTGTCTCACAGCTTCTCTGCGAAG 1001
Oy	308	ThrPheAsnPheAsnLeuThrValGlnThrAlaValProGlyProCysProSerProThr 327
Db	1002	ACTTACCCCTGATTTACTGTTCAGCTTATATACCTGTGCAAGCCAGTAACA 1061
Oy	328	ProSerProSerSerThrThrProSerProAlaSerSerProSerProThrLeuSer 347
Db	1062	CCCATCTCCA-----TCACGCGCCACCCAGCAGTAAGC 1094
Oy	348	ThrProSerProSerLeuMetProThrGlyTyrIysSerMetGluLeuSerAspIleSer 367
Db	1095	ACTGATGATCTTCTTAATTCAGACCCCTCTGCTCCCAATGAGTGGCAGAAACAATCTCT 1154
Oy	368	AsnGluCysArgIysGlyIysPheArgAlaThrIleThrIleValAsp 387
Db	1195	GATGGAGCGTCCCATATTATACAGATATGCACTACACAGCTGTGTATCCACCTGTGTAG 1214
Oy	388	GlyIleLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro 407
Db	1215	GGATCTTCTGAGGTAAATATATCCAGATGACAGCAAGCTCCAGATGACAGAAAGCCAGCT 1274
Oy	408	AspAsnSerLeuMetAspPheIleValThrCysIysGlyAlaThrProThrGluAlaCys 427
Db	1275	GAATATCCATGCTGCGACTTGTGTACTTCCGCAAGGAGTTCCTGACAGATGTCTGT 1334
Oy	428	ThrIleIleSerAspProThrCysGlnIleAlaClnAnArgValCysSerProValAla 447
Db	1335	ACAGCATGTTCTGACCCACAGTGGCGAGGTCTCAGCGCATGTGTATGTATCCCGTGGTC 1394
Oy	448	ValAspGluLeuCysLeuLeuSerValArgArgAlaPheAsnGlySerGlyThrCys 467
Db	1395	GTCACTGATGAATGGCTACTACCATACAGAAAGAGCTTTTGTGAACTGGAGACTGTGT 1454
Oy	468	ValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIle 487
Db	1455	ATMAATATCACTTGGGGGATGACACAGCCAGCCCTTCCGACGCGACTTATTCTGTGA 1514
Oy	488	ProGlyIysAspLeuGlySerProIleArgThrValAsnGlyValIleSerIleGly 507
Db	1515	AATGGA-----GGATCATCTCGCGGACACAAAGGTGTCTCTTCTTCTTGA 1565
Oy	508	CysLeuAlaMetPheValIthrMetValIthrIleLeuLeuTyrLysIleHisLysThrTyr 527
Db	1566	TTCGTCGACGATGTGTGGCCATTGTGGCGCTTTTGTGCTTTTACAGAGATACAAACAATAC 1625
Oy	528	IysProIleGlyIleAsnCysThrArgAsnValValIysGlyIysGlyLeuSerValPheLeu 547
Db	1626	AAACCTATTGAGAGAGTGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1685
Oy	548	SerHisAlaIysAlaProPheSerArgGlyAspArgGluLysAspProLeuLeuGluAsp 567
Db	1686	AGCACTTCAAGACCCCTTTTCTTCCCTAAGACGACTGAGAGAAATCTCTGTGTTGAAAGC 1745
Oy	568	LysPro 569
Db	1746	AMNCA 1751
RESULT 14		
AR018825		
LOCUS		620 bp DNA linear PAT 05-DEC-1998
DEFINITION		Sequence 113 from patent US 5783182.
ACCESSION		AR018825
VERSION		AR018825.1 GI:3973939
REWORDS		unknown.
ORGANISM		unclassified.
REFERENCE		1 (bases 1 to 620)
AUTHORS		Thompson,T.C.
TITLE		Method for identifying metastatic sequences
JOURNAL		Patent: US 5783182-A 113 21-JUL-1998;

AC098114 163541 bp DNA linear HTG 12-JUL-2002
 Rattus norvegicus clone CH230-38P9, *** SEQUENCING IN PROGRESS ***.
 72 unordered pieces.
 AC098114 AC098114.4 GI:21729266
 HTG: HTGS_PHASE1.
 Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
AUTHORS

I. (bases 1 to 162541)

Munzy, D.M., Adams C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Amaratunga, S.L., Amaratunge, H.C., Are, J.R., Ayelle, M., Banks, T., Barabanks, J., Benton, J., Binagke K.C., Blankenburg, K., Bonnin, D., Brachmann, A., Brackley, R., Brown, E., Brown, M., Bryant, N.P., Bowles, S., Brieva-Moore, L., Burkett, C., Burrell, K.L., Byrd, N.C., Bushnell, C., Busch, P., Burnett, C., Carter, M., Cavazos, S.R., Chacko, J.J., Chavez, D., Chen, T.F., Carter, M., Cavazos, S.R., Chowdhry, I., Christopoulos, R., Chung, G., Chen, R., Chen, Z., Coyle, S., Daneshmandi, V., Davidson, C., Cleveland, C., Davison, C., De la Cruz, M., Dedering, D.A., David, R., Delaney, K.R., Delgado, O., Denna, A.L., Ding, Y., Dinh, H.H., Donathwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.J., Feinberg, J., Ferris, J., Ferrara, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Falls, T., Ferraguto, D., Flaggs, N., Garza, N., Garner, T., Garza, N., Gill, R., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Hamilton, K.K., Harris, C., Harris, K., Hart, M., Havlik, J.I., Hawes, A., Hollibaugh, J., Horvath, J., Howard, S., Hughes, J., Huber, J., Huliyil, S., Hummel, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Joliviet, S., Jouhad, S., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratochvil, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, X., Lucier, A., Lucier, R., Martin, R., Martindale, M., Martinez, E., Massey, E., Mashayek, M., Matsumoto, M., Mohabbat, M., Morgan, M., Morris, S., Nguyen, N., Newton, J., Newton, S., Nguyen, N., Nguyen, N., Okwuonu, G., Neall, D., Newton, J., Newton, S., Ngweni, S., Ogih, M., Okwuonu, G., Oraguene, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Petersen, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rieves, M., Rojas, A., Rojokhan, I., Rolfe, M., Ruiz, S.S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Siason, I., Sodergren, E., Sonake, R., Sparks, A., Stanley, E., Tanner, R., Thomsen, K., Taylor, C., Taylor, T., Teiford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Ward-Moore, L., Warren, R., Washington, C., Watlington, S., Wang, S., Ward-Moore, L., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Williams, G., Williams, A., Wu, C., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission
2 (bases 1 to 162541)
Worley, K.C.
2

* NOTE: Estimated insert size may differ from sequence length

(http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)

NOTE: this is a 'working draft' sequence. It currently consists of 72 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record is the result of a preliminary assembly, so some contigs it is available and the accession number will be preserved.

1	1201:	contig of 1201 bp in length
2	1301:	gap of unknown length
1302	2388:	contig of 1087 bp in length
2389	2488:	gap of unknown length
2489	3513:	contig of 1025 bp in length
3514	3613:	gap of unknown length
3614	4731:	contig of 1118 bp in length
4732	4831:	gap of unknown length
4832	5962:	contig of unknown length
5963	5962:	contig of unknown length
5964	7083:	contig of 1121 bp in length
7084	7183:	gap of unknown length
7184	8459:	contig of 1276 bp in length
8460	8559:	gap of unknown length
8560	10098:	contig of 1539 bp in length
10099	10198:	gap of unknown length
10199	11258:	contig of 1060 bp in length
11259	11356:	gap of unknown length
11360	12591:	contig of unknown length
12592	13632:	contig of 1041 bp in length
13633	13732:	gap of unknown length
13733	14876:	contig of 1144 bp in length
14877	14976:	gap of unknown length
14977	16015:	contig of 1039 bp in length
16016	16115:	gap of unknown length
16116	17334:	contig of 1219 bp in length
17335	17434:	gap of unknown length
17435	18682:	contig of unknown length
18683	18682:	contig of unknown length
18684	20046:	contig of 1184 bp in length
20047	20146:	gap of unknown length
20147	21488:	contig of 1342 bp in length
21489	22588:	gap of unknown length
22589	22715:	contig of 1127 bp in length
22716	22815:	gap of unknown length
22816	23923:	contig of 1110 bp in length
23924	24942:	gap of unknown length
24943	25537:	gap of unknown length
25537	25636:	gap of unknown length
25637	26722:	contig of 1086 bp in length
26723	26832:	gap of unknown length
26833	27978:	contig of 1156 bp in length
27979	28078:	gap of unknown length
28079	29407:	contig of 1329 bp in length
29408	29507:	gap of unknown length
29508	30541:	contig of 1034 bp in length
30542	30641:	gap of unknown length
30643	31988:	contig of unknown length
31989	31988:	contig of unknown length
31990	33108:	contig of 1110 bp in length
33109	33408:	gap of unknown length
33409	34411:	contig of 1203 bp in length
34412	34511:	gap of unknown length
34512	36645:	contig of 2134 bp in length
36646	36745:	gap of unknown length
36746	38188:	contig of 1443 bp in length
38189	38288:	gap of unknown length
38289	39579:	contig of unknown length
39580	41328:	contig of 1749 bp in length
41329	41428:	gap of unknown length
41429	42540:	contig of 1112 bp in length
42541	42640:	gap of unknown length

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result	No.	Score	Query	Length	DB	ID	Description
	1	3040	99.3	2303	19	AAT96034	Rat kidney injury
	2	2700	88.2	2213	21	AAAT96729	Polynucleotide iso
	3	21f3	70.7	2659	18	AAT69321	Murine metastatic
	4	21c3	70.7	2659	18	AAT69328	Murine metastatic
	5	21f3	70.7	2659	18	AAT69328	Human NHE CDNA seq
	6	21f3	70.7	2659	18	AAT69328	Human NHE CDNA seq
	7	21e2	70.6	2683	24	ABK33556	Human NHE CDNA seq
	8	941	30.7	620	18	AAT69340	Murine metastatic
c	9	729	23.8	524	18	AAT69342	Murine metastatic
	10	717	23.4	473	18	AAT69342	Murine metastatic
	11	627.5	20.5	2115	16	AAO96055	Sequence encoding
	12	627.5	20.5	2115	16	AAO96055	Human melanoma ant
	13	627.5	20.5	2115	16	AAO96055	Human melanoma ant
	14	627.5	20.5	2115	16	AAO96055	Human melanoma ant
	15	627	20.5	2172	22	AAK55525	Human melanoma ant
	16	622.5	20.3	1986	22	AAO07346	Modified tumour-as
	17	622.5	20.3	1986	22	AAO22098	Human gp100m nucle
	18	622.5	20.3	1986	22	AAO22098	Human gp100m nucle
	19	622.5	20.3	2534	22	AAH20120	Plasmid C5H6p100M
	20	622	20.3	2534	22	AAH20120	Plasmid C5H6p100M
	21	521	20.3	2131	18	AAT96760	pHE17 CDNA. Homo
	22	513.5	20.0	1679	13	AAT96760	Melanoma-specific
	23	508.5	20.0	1679	13	AAT96760	Human osteoblast d
	24	524	17.1	335	16	AAO93186	Human osteoblast d
	25	524	17.1	335	16	AAT69350	Murine metastatic
	26	524	17.1	335	18	AAT69319	Murine metastatic
	27	524	17.1	335	18	AAT69321	Murine metastatic
	28	504	16.5	354	18	AAT69323	Murine metastatic
c	29	471	15.4	286	18	AAT69348	Murine metastatic
c	30	462.3	15.1	318	22	ABM74956	Murine metastatic
c	31	462.3	15.1	318	22	ABM74956	Murine metastatic
c	32	462.3	15.1	318	22	ABM74956	Murine metastatic
c	33	462.3	15.1	318	22	ABM74956	Murine metastatic
c	34	462.3	15.1	318	22	ABM74956	Murine metastatic
c	35	462.3	15.1	318	22	ABM74956	Murine metastatic
c	36	462.3	15.1	318	22	ABM74956	Murine metastatic
c	37	462.3	15.1	318	22	ABM74956	Murine metastatic
c	38	462.3	15.1	318	22	ABM74956	Murine metastatic
c	39	462.3	15.1	318	22	ABM74956	Murine metastatic
c	40	418	13.7	309	18	AAT69333	Murine metastatic
c	41	409	13.4	260	18	AAT69346	Human breast cell
c	42	405.5	13.2	478	22	ABA24257	Human breast cell
c	43	405.5	13.2	478	22	ABA23280	Human foetal liver
c	44	405.5	13.2	478	22	ABA23054	Human foetal liver
c	45	405.5	13.2	478	22	AAK01537	Probe #1520 for gels
							Human brain expres

ALIGNMENTS

RESULT 1	
AAAT96034	
ID	AAAT96034 standard; cDNA; 2303 BP.
AC	
CD	AAAT96034;
DP	
DT	21-MAY-1998 (first entry)
DS	
DE	Rat kidney injury related molecule (KIM) cDNA clone 4-7.
DD	
XX	Kidney injury related molecule; KIM; rat; renal disease; injury;
KK	nephritis; tissue regeneration; therapy; as.
KK	
XX	Rattus sp.
XX	
XX	Location/Qualifiers
PH	Key
PH	107.1825
CDS	/*tag= a
FT	


```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 623 TTGACGCTGGCCCTCAGGTGAAAGTACTGCTTTCCGAGATACGCGCGGCATAC 682
Oy 221 TLePcOtLeVasVAlVasVAlVasVAlVasVAlVasVAlVasVAlVasVAlVasVAl 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 683 ATTCCATCTCGAAGTGAAGGTGTGTGTATGATACAGATCAGATCTGCTGATTCGTG 742
Oy 241 ThrMetTyGlnVasAsnAraGAsnSerSerAspGluThrPheLeuAraGAsnLeuPro 260
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 743 ACATATGCCAGAGATGATGACAGACTGCTGTGATGATCTTCTCCAGACCTCCCC 802
Oy 261 TLePheAspVAlVAlVAlVAlVAlVAlVAlVAlVAlVAlVAlVAlVAlVAlVAl 280
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 803 ATGCTGCTCGATGCTCTCATTCATGATCCAGCCACTTCTCAAGGACTCTGCCATTCC 862
Oy 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisLeuAsn 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 863 TACAAGTGGACCTTGTGGGACAACATGGCTGCTTGTCTCCACAATCACAATTTGAAT 922
Oy 301 HisThrTyValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaVal 320
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 923 CACATTTATGTGTCTCAATGAACTTCAACCTTAACTCAAGCTGCAAGTGCAGTGCC 982
Oy 321 GlyProCysProSerProThrProSerProSerProSerProSerProSerProSerPro 338
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 983 GGCCATGCTCCCTCCCTCGCTTCTGACCTGCTCCCTCACTCAACTCGCCCTCACT 1042
Oy 339 AlaSerSerProSerProThrThrLeuSerThrProSerProSerLeuMetProThrGly 358
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1043 CGCCCTCACTCTGCCCATTTATCAACACCTTAGCCCTCTTAAATGCCCTAGTGTTAC 1102
Oy 359 LysSerMetGluLeuSerAspLeuSerAsnGluAsnGlyArgLeuAsnArgTyrGlyTyr 378
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1103 AATCCATGGAGCTGTGAGTACATTCATGAAGACGCGGATACAGATATGGCTAC 1162
Oy 379 PheCThrThrThrLeuThrValAspGlyLeuLeuGluValAsnThrIleGlnValAla 398
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1163 TTCAGACCCACATCACAATTTGAGAGGGATCTCGAAGTCAGCATCATGCAGATAGCA 1222
Oy 399 AspValProLeuProThrLeuGlnProAsnSerLeuMetAspPheLeuValThrCys 418
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1223 GATGTCCTCCATGCCACGCGAGCTGCCACCTCTGTGAGCTTCACTGTGAGCTGC 1282
Oy 419 LysGlyAlaThrProThrGluAlaCysThrIleLeuSerAspProThrCysGlnIleAla 438
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1283 AAGAGGCGCACCCCATCysProValAlaValAspGluValCysLeuValAsnArg 458
Oy 439 GlnGluThrValCysProValAlaValAspGluValCysLeuValAsnArg 478
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1343 CAGAACCGGGTGTGACGCGCTCTGCTGTGATGGCTGCTGCTGTCTGTGAAGA 1402
Oy 459 AlaPheAsnGlySerGlyThrTyCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1403 GCTTCAATGGGTGTGGCACTTCTGTGATTTCACTCTGGGAGATGATGACGGCTG 1462
Oy 479 AlaLeuThrSerAlaLeuThrSerLeuProGlyLysAspLeuGlySerProLeuArgThr 498
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1463 GCCTTCACGACGACCTGTATCTATCTCCCTGCGCAAGACCCAGACTCCCTCTGAGACA 1522
Oy 499 ValAsnGlyValLeuThrLeuSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1523 GGAATGGTGTCTGATCTCATATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1582
Oy 519 LeuLeuThrLysLeuHisLysThrTyCysValAsnPheThrLeuGlyAsnValVal 538
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1583 TTGCTGTGTAACAAAACACAGAGCGTACAGACCATATGAGAACTGGCCCGACGACGCGTC 1642
Oy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAsp 558
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1643 AGGGCAAGAGCGCTGAGGTCTCTCTCAAGCCAGCGAAGCCCGCTTCTCCAGAGAGAC 1702
Oy 559 ArgGlnLysAspProLeuLeuGlnAspLysPro 569
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

Db 1703 CAGGAGAAGGATCCATTGCTCCAGGACAAGCCA 1735
RSJUT 3
XNT69318
ID AAT69318 standard; cDNA, 2669 BP.
AC AAT69318;
DT 26-FEB-1998 (first entry)
XX Murine metastatic nucleic acid sequence.
XX Mouse; murine; tumour; cancer; metastatic sequence; detection:
KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;
KW hypertrophy; screening; ss.
XX Mus musculus.
XX X09718454-AZ.
XX W09718454-AZ.
PD 22-MAY-1997.
XX 15-NOV-1996; 96WO-US18567.
PR 30-JAN-1996; 96US-0594031.
PR 16-NOV-1995; 96US-0068838.
PA (THOM)/ THOMPSON T.
PI Thompson T.
XX WPI; 1997-289397/26..
XX Identifying tumour metastatic sequences - by introducing transfected
XX cells into host animals and analysing primary and metastatic
PT sequences by differential display PCR
XX Disclosure: Fig 12BT; 102pp; English.
XX Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old
XX mice and the cells were isolated from the capsule of mice.
XX Cells were isolated under the renal capsule of mice.
XX Reconstructions were harvested 5 weeks later, when they showed
XX signs of distress from the tumour burden. Metastatised tumours were
XX isolated from a site outside the renal capsule. RNA was isolated
XX from primary tumours and metastases, reverse transcribed and
XX subjected to differential display PCR. The sequences were analysed
XX and the results were used to detect, diagnose and treat disorders related
XX to metastasis, or treat malignant or non-malignant disorders, e.g.
XX hyperplasia, dysplasia and hypertrophy. The metastatic sequence can
XX be used to screen a biological sample for metastasis, and it or its
XX expression product may also be used to treat a metastatic disorder.
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;
Alignment Scores:
Pred. No.: 1,61e-155 Length: 2669
Score: 2163.00 Matches: 394
Percent Similarity: 82.57% Conservative: 75
Best Local Similarity: 69.37% Mismatches: 85
Query Watch: 10.66% Indels: 14
Dbs: 1 Gaps: 1
US-09-943-075A-2 (1-572) x AAT69318 (1-2669)
Oy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20
||||| ||| ||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 92 ATGGAATCTCTCTACTATTTCTGGGATTTCTCTGCTGCTCAAGATTGCACTGAT 151
Oy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnThrProAspHisMetArgGlu 40
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 152 GCCGCCAAGACGATTTTCATGCTGCTGGCAATGAAGACCTTCTGCTTACATAGGGAG 211

```


CC Mouse urogenital sinus (UGS) tissue was isolated from 17 day old
 CC mouse embryos. The UGS cells were infected with retroviruses,
 CC and the cells were grown in 96 well plates. After 5 days, the
 CC reconstitutions were harvested 5 weeks later. The cells showed
 CC signs of distress from the tumour burden. Metastatised tumours were
 CC isolated from a site outside the renal capsule. RNA was isolated
 CC from primary tumours and metastases, reverse transcribed and
 CC subjected to differential display PCR. The sequences were analysed
 CC to obtain metastatic sequences, e.g. the present sequence. The
 CC method can be used to detect, diagnose and treat disorders related
 CC to hyperplasia, dysplasia and hyperproliferation. The metastatic
 CC expression product may also be used to treat a metastatic disorder.
 XX

Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;

Alignment Scores:

Pred. No.:	1,616-155	Length:	2669
Score:	2163.00	Matches:	394
Percent Similarity:	82.57%	Conservative:	75
Best Local Similarity:	69.37%	Mismatches:	85
Query Match:	70.66%	Indels:	14
	18	Gaps:	1

US-09-943-075A-2 (1-572) x AT69328 (1-2669)

Oy 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20
 Db 92 ATGAATGCTCTACTATTCTCGGATTTCTGCTCGCTGCGAAGATTGCCACTTGAT 151
 Oy 21 AlaAlaValArgPheArgPheArgPheArgPheArgPheArgPheArgPheArgPhe 40
 Db 152 GCGGCCAACCAATTCATGATCTCTGGCAATGAAAGACCTTCGCTTACATGGGAG 211
 Oy 41 AsnAsnInLeuArgGlyTrpSerAspGluAsnGlyTrpAspGluGlnLeuTrp 60
 Db 212 CACATCAATTAAGTGGCTGCTCTGATGAAATGCTGAATGAAATGCTACCCA 271
 Oy 61 ValTrpArgAsnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 80
 Db 272 GTGTGGAGCGGGAGACATGAGTGGAAAACTCTGGAGGGAGGCGCTCTGAGCG 331
 Oy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
 Db 332 GTCTGACACGGATCTACACCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
 Oy 101 ValPheProArgCysCysGlyLeuAspAlaAsnGlyValSerIleValTrpGlyGlyAsnGly 120
 Db 392 ATATTCCTTAGATGCCAAAGCAAGATGCCAATGCCAATAGTCTATGAGAAGATGC 451
 Oy 121 ArgSerAspLeuGluLeuAlaSerAspProValTrpValTrpValTrpValTrpValTrp 140
 Db 452 AGAATGAGGCTGGTATCTCTGCTGATCCATATGTTTACATGCAGCAGCATGCTGAG 511
 Oy 141 AspGluAspTrpGlyLeuAspAsnThrSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 160
 Db 512 GACAGTACGGGAAATGCCACGGCCAAAGCCATCATCAACCTTCTCCCTCATGGGAAA 571
 Oy 161 ProPheProArgProHisIleGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 180
 Db 572 CTTCTTCTCCACGCCCGGATGGAGAGATGGATTTCTTCTTCTTCTTCTTCTTCTTCT 631
 Oy 181 Gly 200
 Db 632 GTCATGATTTCCAGAAATTTGGACGATGTTTCTGAGAGATTTCTGTGACACAGCCAA 691
 Oy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgGlyIleGlyArgAlaTrp 220
 Db 692 GTGACATCTGGGCTCTACTCTGAGAGTGGTGTCTACAGAGATGGCGGCGATAT 751
 Oy 221 IleProIleValLeuValValValValValValValValValValValValValValVal 240

Db 752 GTTCCCATCCCAAGTGAAGATGTGTACGTGGTAACAGATCAGATCTCTGTGTGTG 811
 Oy 241 ThrMetTrpGlyGlyAsnAspArgAsnSerSerAspGlyLeuThrPheLeuArgAspLeuPro 260
 Db 812 ACTATGTTCCAGAGAACGTCAGAAATTCATCCGAGAACCTTCTCTCAAGATCTCCCT 871
 Oy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTrpSerAlaIleSer 280
 Db 872 ATATGTTGTGTGCTGATGATGATCTAGCCACTTCTCTCAATTAATTTCTACCATTAAC 931
 Oy 281 TrpTrpTrpAspPheGlyAspAsnThrGlyLeuPheValSerAsnIleThrLeuAsn 300
 Db 932 TACAGCTGGAGCTTGGGGATTAATCTGGCTCTGTTTGTTCACCAACATCATCTGTGAT 320
 Oy 301 HisThrTrpValIleAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 Db 992 CACACATGTGCTCTCAATAGAACCTTCAGCCTTAACCTCTGTAAGACGACGACACA 1051
 Oy 321 GlyProCysProSerProSerProSerProSerProSerProSerProSerProSerPro 340
 Db 1052 GGACCTTGTCCGCCCGCCACCCACCCAGCCAGCCT - - - - - 1087
 Oy 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTrpLysSer 360
 Db 1088 - - - - - PCAAACCCACCCCTCTCTTAGGACCTTCCTGGTGGACAAACCC 1129
 Oy 361 MetGluLeuSerAspIleSerAsnGlyAsnCysArgIleAsnArgTrpGlyTrpPheArg 380
 Db 1130 CTGGAGCTGAGTAGGATTCCTGTATGAAATGCCAGATTTACAGATATGGCCTTTCAA 1189
 Oy 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 1400
 Db 1190 GCACCATCTCAATTTGAGAGGATCTTGAAGCTTACATCATCCAGATGACGACGCTC 1249
 Oy 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
 Db 1250 CTGATGCGCGTCCCATGGCTGAAAGCTCCCTTAATAGACTTGTCTGACCTGCCAGGG 1309
 Oy 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 1440
 Db 1310 AGCATTCACAGAGGCTGTACATCATTTCTGACCCACCTCCGAGATCACCCAGAAC 1369
 Oy 441 ArgValCysSerProValAlaValAspGlyLeuLeuSerValArgArgAlaPhe 1460
 Db 1370 ACAGTCTGCAGCCCTGTGATGTGATGATGATGATGATGATGATGATGATGATGATG 1429
 Oy 461 AsnGlySerGlyThrTrpCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
 Db 1430 NATGGTCTGGGAGTACTGTGTGAACTCACCTCCGGGGATGACACAGCTGGCTCTC 1489
 Oy 481 ThrSerAlaLeuSerIleSerProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
 Db 1490 ACAGAGACCCCTGATGCTGTCTCCACAGACCCAGCCAGCTCCCTTTTAAAGATGGCAAC 1549
 Oy 501 GlyValLeuLeuSerIleSerIleLeuAlaMetPheValThrMetValThrIleLeuLeu 520
 Db 1550 AGTGGCCTGATCTCCGCTGGCTCTGGGCAATTTTGTCTGCTGATGATCTCTCTTTGGTG 540
 Oy 521 TrpTrpLysHisLysThrTrpLysProIleGlyAsnCysThrArgAsnValValLysGly 540
 Db 1610 TACAAAMACACAGAGTAAACCAACCAATGAAATAGTCTCGGAATGTGTGAGAAGC 1669
 Oy 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
 Db 1670 AAAGGCTGAGTGTCTTCTCAACCGTGCACCGTGTCTTCTCCGGGAACACAGGAA 1729
 Oy 561 LysAspProLeuLeuGlnAspLys 568
 Db 1730 AAGGATCCGCTACTCAAAAGCAA 1753
 RESULT 5
 AAC55715
 ID AAC55715 standard; cDNA; 2669 BP.

XX AAC55715;
 AC 17-JAN-2001 (first entry)
 AT Human NMB cDNA sequence from Genbank X76534.
 DE Human; differentially regulated gene; macrophage development;
 KW diagnosis; matrix metalloproteinase 19; MMP19; antiarthritic;
 KW antiinflammatory; destructive macrophage development inhibitor;
 KW arthritis; colorectal cancer; immune response; ss.
 OS homo sapiens.
 PN WO20005373-A2.
 XX 21-SEP-2000.
 PD 15-MAR-2000; 2000WO-US06883.
 XX 15-MAR-1999; 99US-0124530.
 PR (EOSB) EOS BIOTECHNOLOGY INC.
 PA Murray R.
 PI WPI: 2000-628200/60.
 PD
 XX Screening drug candidates comprises adding a drug to a cell expressing
 PT an expression profile gene and determining the effect of the drug on
 PT the expression of the expression profile gene
 XX
 PS Claim 1: Page -: 99pb: English.
 CC The present invention describes a method for screening drug candidates.
 CC The method comprises adding a drug to a cell that expresses an
 CC expression profile gene encoding a protein encoded by 5 sequences of
 CC defined base pairs as given in C55638, C55642, C55644 and
 CC C55653 or a sequence represented by Genbank accession number X92521,
 CC X62466, J04130, X62087 and X76534 (or a fragment) and determining the
 CC effect of the drug on the expression of the expression profile gene.
 CC A method of screening drug candidates comprising providing an
 CC antibody, is useful for treating destructive macrophage disorders (DMD)
 CC by inhibiting DMD development in a cell of an individual having
 CC arthritis. Antibodies to MMP-19 are useful for localising a therapeutic
 CC moiety preferably cytotoxic agent or a radioisotope to colorectal cancer
 CC tissue. A composition comprising MMP-19 is useful for eliciting an
 CC immune response in an individual. C55635 to C55710 represent human
 CC cDNA sequences that regulate genes of the invention. The present sequence
 CC represents the NMB cDNA sequence according to the Genbank
 CC accession number X76534.
 CC N.B. The present sequence is not given in the present specification,
 CC but it is specifically claimed by its Genbank accession number.
 XX
 90 Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,616-155 Length: 2669
 Score: 2163.00 Matches: 394
 Percent Similarity: 82.57% Conservative: 75
 Best Local Similarity: 69.37% Mismatches: 85
 Query Match: 70.66% Indels: 14
 DB: 21 Gaps: 1

US-09-943-075A-2 (1-572) x AAC55715 (1-2669)

Oy 1 MetGluSerLeuGlyValPheLeuLeuAlaAGlyLeuProLeuGln 20
 Db 92 ATGGAAATGCTCTACTATTTCTGGGATTTCTGCTCGCTGCAAGATTGCCACTTGAT 151
 Oy 21 AlaAlaValArgPheArgAspValIleGlyHisGluGlnTyrProAspHisMetArgGlu 40
 Db 152 GCSCCAACCAATTGATGCTCGTGGCAATGAAGACCTTCTCTACTATGAGGAG 211

Oy 41 AsrAsnGlnLeuArgGlyTrpSerSerAspGluAndLutPspAspGluGlnLeuTyrPro 60
 Db 212 CACAATCAATTAATGGCTGGTCTTCATGAAATGACTGGATGAAATCACTACCCA 271
 Oy 61 ValTrpArgArgGlyGluGlyValGlyTrpGlySerPheGluGlyValArgValGluAla 80
 Db 272 GTGTGAAGCGGGAGACATGAGTGGGAAAAAATCTCTGGAGGGAGCGGTGTCGACGC 331
 Oy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
 Db 332 GTCTCGACCTAGTGATGCTCACCAGCCCTCGTGGCTCAATTAACATTTGGCGTGAACCTG 391
 Oy 101 ValPheProArgCysGlnGlyGluAspAlaAsnGlyValGlyValTrpGluArgAsnCys 120
 Db 392 ATATTCTCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAGAAGAACTGC 451
 Oy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrGlyValLeuAsp 140
 Db 452 AGAATGAGAGCGGTGTTTCTGCTGCTCATATGTTTACACATGGACACATGGCTCGACG 511
 Oy 141 AspTyrLeuAspTrpGluAspAsnThrSerGlnGlyGluHisLeuArgPheProAspGlyLys 160
 Db 512 GACGTGACCGGGGAAATGCCACCGCCAAAGCCATCATAACTCTCTCCCTGATGGGAAA 571
 Oy 161 ProPheProArgProHisGlyArgLysLysLysLysLysLysLysLysLysLysLysLys 180
 Db 572 CTTTTCCTTCAACACCCCGGATGGAGAGATGATATTCATCTAGCTCTTCCACACGCTT 631
 Oy 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
 Db 632 GGTGAGTATTTCCAGAATTTGGGACGATGTTCAGTGAGAGATTTCTGTGAACACACAGCCAT 691
 Oy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220
 Db 692 GTGACATTTGGGCTGACTGCTGGATGCTGCTACAGAGACATGGACGGGCTAT 751
 Oy 221 IleProIleSerLysLysValAspValTyrValIleThrAspGlnIleProIlePheVal 240
 Db 752 GTTCCCATCGCAAGTGAAGATGTTAGCTGGTAAACAGATCAGATTCCTGTTGTG 811
 Oy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspLysLysLysLysLysLysLysLys 260
 Db 812 ACTATGTTCCAGACAGACCATCGAAATTCATCCAGCAAACTCTCTCAAGATCTCCCC 871
 Oy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
 Db 872 ATTATGTTTGAATGCTGCTGATGATGATGCTGCACTCTCAATATTCATACCAATTAAC 931
 Oy 281 TyrTrpArgPheGlyGlyAsnThrGlyLeuPheValSerAsnAsnHisPheLeuAsn 300
 Db 932 TACAAGTGGAGCTCGGGGATAATAGTGGCTGTTTGTGTTTCCACCAATCATCTATGTAAT 991
 Oy 301 HisThrTyrValIleAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 Db 992 CACACCTATGCTCAATGGAACTTTCAGCTTAACTGCTGCTGAAAGCTGACGACCA 1051
 Oy 321 GlyCysCysProSerProThrTyrProSerProSerSerSerSerSerProSerProAlaSer 340
 Db 1052 GGACCTTGTCTCCGCCACCCACCCACCCACCCACCCACCT----- 1087
 Oy 341 SerProSerProThrLeuSerThrProSerThrProSerLeuMetProThrGlyTyrLysSer 360
 Db 1088 -----TCAAAACCCACCCCTCTTTTAGGACTGCTGGTGACACCCC 1129
 Oy 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380
 Db 1130 CTTGAGCTGAGTAGGATGCTGCTGATGAAATGCTGCAAGATTAAACAGATATGCCCATTTCAA 1189
 Oy 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleIleIleIleIleIleIleIle 400
 Db 1190 GCCACCTTCACATNTGTAGGGGATCTTAGAGGTTAAACATCATCATCATCATCATCATCATCAT 1249

QY 401 ProfileProTherLeuInProAspSenSerLeuMetAspPheIleValThrCysGly 420
 DB 1250 CTGATCGCGGCGCATCGGCTCGAAGCTCCTTAATAGACTTTGTGTGACCTGCCAAGG 1309
 QY 421 AlaThrProThrGluAlaCysPheThrIleSerAspProThrCysGlnIleAlaGln 440
 DB 1310 AGCATCCCGAGGAGTGTGTACCATCATTTCTGACCCACCTCGGAGATCCACCAAC 1369
 QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuSerValArgAlaPhe 460
 DB 1370 ACGATCGACGCGCTGTGGATGGATGATGTGTCTGTGACTGTGACGAGCACTTC 1429
 QY 461 AsnGlySerGlyThrTyrrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
 DB 1430 AATGGTCTGGGACGTACTGTCTCAACCTCAACCTCGGGGATGACACAAAGCTGGCTCTC 1489
 QY 481 ThrSerAlaLeuIleSerIleProGlyAspLeuGlySerProLeuArgThrValAsn 500
 DB 1490 AGAGACCCCTGATTTCTGTCTCGACAGACCCAGCTCGCTTAAAGATGGCAAC 1549
 QY 501 GlyValLeuIleSerIleGlyCysAlaMetPheValThrMetValThrIleLeu 520
 DB 1550 AGTGCCTGATCTCGCTGTGTGTGGCCATATTTGCTGATCTGATCTGCTGTGTG 1609
 QY 521 TyrIlySerHisIlyThrTyrrIleProIleGlyAsnCysThrArgAsnValValGly 540
 DB 1610 TACAAACACAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1669
 QY 541 LysGlyIleSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArg 560
 DB 1670 AAGGCTCGAGTGTCTTCTCAACCTCGCAAGCGGTGTCTTCCCGGAACCAAGAA 1729
 QY 561 LysAspProLeuLeuInAspLys 568
 DB 1730 AAGGATCCCTACTCAAAACCA 1753

RESULT 6
 ID ABQ88185 standard; cdna; 2669 BP.
 AC
 AC ABQ88185;
 DT 18-SEP-2002 (first entry)
 XX Human osteoblast differentiation related cdna SBO ID NO 92.
 DE Human; osteoblast; stem cell differentiation; bone tissue deposition;
 KW osteocalcin, osteopontin, osteopontin, etc.
 QS Homo sapiens.
 PN WO20020301-A2.
 PD 27-JUN-2002.
 PF 18-DEC-2001, 2001NOV-USA8276.
 PR 18-DEC-2000, 2000US-255882P.
 PR 24-APR-2001, 2001US-285591P.
 XX (GENE-) GENE LOGIC INC.
 PA (PROC) PROCTER & GAMBLE CO.
 PI J. D. Axelrod, Cook JS, Jaiswal N, Einstein R, Houghton A;
 PI Mett L;
 DR WPI; 2002-557663/59.
 XX Use of genes and their expression profiles associated with osteoblast
 differentiation for screening modulators bone formation, for diagnosing
 or treating e.g. osteoporosis, or as markers for the differentiation
 process

PS Claim 1: SEQ ID NO 92: 78bp + Sequence Listing; English.
 XX The invention relates to genes and their expression profiles are used
 CC for: screening modulators of precursor stem cell differentiation into
 CC osteoblasts or bone tissue deposition.
 CC (a) diagnosing abnormal deposition of bone tissue, abnormal rate of
 CC osteoblast formation or osteoporosis; or
 CC (c) treating or monitoring treatment of the conditions cited in (b), or
 CC monitoring the progression of bone tissue deposition.
 CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
 CC osteoporosis, alcohol-induced osteoporosis, osteoporosis associated with
 CC osteoporosis, osteoporosis associated with bone formation or bone loss conditions
 CC that involve altered bone metabolism (e.g. idiopathic juvenile
 CC osteoporosis), skeletal bone metabolism linked to breast cancer, mastocytosis,
 CC Fanconi syndrome or fibrous dysplasia. The present sequence is that of an
 CC osteoblast differentiation associated cdna marker of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at tip.wipo.int/pub/published_pat_sequences.

XX Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 other;

Alignment Scores:

Pred. No.: 1,61e-155 Length: 2669
 Score: 236.0 Macchi: 2669
 Best Similarity: 83.5% Clustal: 75
 Best Local Similarity: 69.37% Mismatches: 85
 Query Match: 70.66% Indels: 14
 DB: 24 Gaps: 1

US-09-943-075a-2 (1-572) x ABQ88185 (1-2669)

QY 1 MetGluSerCysGlyValLeuValPheLeuLeuAlaIleGlyLeuProLeuGln 20
 DB 92 ATGGAATGTCTCTACTACTATTTCTCTGGGATTTCTCTGCTGCAAGATTCGCACTTGTAT 151
 QY 21 AlaAlaIleArgPheArgAspValLeuGlyHisGluInTyrrProAspHisMetArgGlu 40
 DB 152 GCCCCCAACGATTCATGCTGTGGCATGAAACCTTCCTGTCTACATGAGGAG 211
 QY 41 AsnAsnGlnLeuArgGlyTyrrSerSerAspGluAsnGluTrpAspGluGluLeuTyrrPro 60
 DB 212 CACAAATCAATTAATGTGTCTCTCTGATGAAATGACTGGAAATGAAACCTTACCCA 271
 QY 61 ValTrpArgArgGlyGluGlyArgTyrrPheAspSerTrpGluGlyArgValGluAla 80
 DB 272 GTGTGGAGCGGGAGACATCAGTGGANNAATCTCTGGAAGGAGGCGGTGCGAGCG 331
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValAlaAsnLeu 100
 DB 332 GTCTTGACCAAGTACCTCACCAGCCCTCTGGCTCAATATACATTTGCGGTGAACCTG 391
 QY 101 ValPheProArgGlyGlnGlyGluAspAlaAsnGlyAsnIleValTyrrGluAspGlyCys 120
 DB 392 ATATTCCTAGATCCAAAGGAAGATGCGATCCATCTATGAGAGAACTGC 451
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrrValTyrrAsnTrpThrGlyAlaAsp 140
 DB 452 AGAATAGAGCTGTGTATCTGTCTGATCCATATGTTTACACTGGACACATGCTCGAG 511
 QY 141 AspGluAspTrpGluAspThrSerGlnGlyGlnHisLeuArgPheArgPheAspGlyCys 160
 DB 512 GACAGTGACGGGAAATGGCACCGCCCAACCATCATACCTCTCTCTGATGGGAA 571
 QY 161 ProPheProArgProHisGlyArgLysTyrrAsnPheValTyrrValPheHisThrLeu 180
 DB 572 CTTTTCCTCACCACCCCGGAGAGATGGAATTCATCTACCTCTCTGACACACTT 631
 QY 181 GlyGlyTyrrPheGlnGlyGlyGlnGlySerAlaArgValSerIleAsnThrValAsn 200
 DB 632 GGTCACTATTCAGAAATGGAGATGTTCTGAGAGAGTTTCTGTGACACAGCAAT 691

KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;
 KW hypertrophy; screening; ss.
 XX Mus musculus.
 OS WO9718454-A2.
 PN 22-MAY-1997.
 XX 22-MAY-1997.
 XX 15-NOV-1996; 96WO-US18567.
 XX 30-JAN-1996; 96US-0594031.
 PR 16-NOV-1995; 95US-0006838.
 XX (THOM) THOMPSON T.
 XX Thompson T.
 PI WPI: 1997-289397/26.
 DR Identifying tumour metastatic sequences - by introducing transfected
 XX mouse embryos. The UGS cells were infected with retroviruses,
 XX cultured and implanted under the renal capsule of mice.
 XX Reconstructions were harvested 5 weeks later, when they showed
 CC signs of distress from the tumour burden. Metastatised tumours were
 CC isolated from a site outside the renal capsule. RNA was isolated
 CC from primary tumours and metastases, reverse transcribed and
 CC subjected to differential display PCR. The sequence transes were analysed
 CC to obtain metastatic sequences (e.g. the present sequence). Related
 CC sequences were used to screen a biological sample for metastasis, e.g.
 CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can
 CC be used to screen a biological sample for metastasis, and it or its
 CC expression product may also be used to treat a metastatic disorder.
 XX Sequence 620 BP; 143 A; 207 C; 128 G; 142 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 5,05e-63 Length: 620
 Score: 941.00 Matches: 183
 Percent Similarity: 91.75% Conservative: 6
 Best Local Similarity: 86.83% Mismatches: 35
 Identity Match: 86.74% Gaps: 1
 DB: 18

US-09-943-075a-2 (1-572) x AAT69340 (1-620)

OY 274 LeuAnThySerAlaTleSerThyTrpAsnPhedGlyAspAsnThrGlyLeuPheVal 293
 DB 2 CAGACACTTCCCATTTCTCCANAGGGAGCTTGGGGACACACTGGCTGTTC 61
 OY 294 SerAsnAnshThrLeuAnshThrTyValLeuAnGlyThrPheAsnPheAsnLeu 313
 DB 6 TCCACACATACACTTGAATACACACTTATGTGCTCAATGGACCTTCAACCTTAACCTC 121
 OY 314 ThxValGlnAlaValPcGlyPcGlyPcSerProThrPhePcSer-----ProSer 331
 DB 122 ACCGTGCAAACTGCAGTGGCGGGGCATCCCTCCCTTCGACTTCGCTGCCA 181
 OY 332 SerSerThrSerProAlaSerProSerProSerProThrLeuSerThrSerPro 351
 DB 182 CTTTCAACTCCCGCCACTCCGCCCTCCTGCGCCATATTCACACACTAGCCGC 241
 OY 352 SerLeuSerProThrGlyThrValSerMetGluLeuSerAspLeuSerAsnGluSer 371
 DB 242 TCTTTAATGCTACTGGTTACAAATCCATGGAGCTGAGTGACATTCATGAATGCTC 301

OY 372 AraTleAsnArgTyGlyThrPheAsnGlnAThrIlePThrIleValAspGlyLeuLeuGlu 391
 DB 302 CGAATAACAGATATGCTACTTTCAGAGCCACCATCACAAATGTAGAGGGGATCTCGAA 361
 OY 392 ValAsnIleIleGlnValAlaAspValProIlePThrLeuGlnProAspAsnSerLeu 411
 DB 362 GTCACTATCATCATGATAGCATGCCATCCCATGCCACACCCGACCTGCCAATCTCCG 421
 OY 412 MetAspPheIleValThxCysTyAsGlyValAThrProThrGluAlaCysThrIleTleSer 431
 DB 422 ATGACTTCACTGTGCTGCAAGAGGGGCCACCCCATGGAAGCTGTACGATCATCTCC 481
 OY 432 AspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeu 451
 DB 482 GACGACCTTCCAGATCCGCCAGACCCGGTCTCCAGCCCTGTGGCTGTGATGGGCTG 541
 OY 452 CysLeuLeuSerValArgAlaPheAsnGlySerGlyThrTyCysValAsnPheThr 471
 DB 542 TCGCTGCTGTGAGAGAGAGCTTCAATGGGTGTGGCACCTACTGTGTGAATTCAC 600
 OY 472 LeuGlyAspAspAlaSer 477
 DB 601 CTGGGAGATGATCAAGC 618
 RESULT 9
 AAT69332/c
 ID AAT69332 standard; cDNA; 524 BP.
 XX AAT69332.
 XX AAT69332.
 DT 26-FEB-1998 (first entry)
 DE Murine metastatic nucleic acid sequence.
 XX Mouse; murine; tumour; cancer; metastatic sequences; detection;
 KW diagnosis; treatment; metastasis; hyperplasia; dysplasia;
 KW hypertrophy; screening; ss.
 OS Mus musculus.
 XX WO9718454-A2.
 XX 22-MAY-1997.
 XX 15-NOV-1996; 96WO-US18567.
 PR 30-JAN-1996; 96US-0594031.
 PR 16-NOV-1995; 95US-0006838.
 XX (THOM) THOMPSON T.
 XX Thompson T.
 PI WPI: 1997-289397/26.
 DR Identifying tumour metastatic sequences - by introducing transfected
 XX cells into host mammal and analysing primary and metastatic
 XX sequences by differential display PCR
 XX disclosure; Fig 12CV: 102pp; English.
 XX Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old
 XX mouse embryos. The UGS cells were infected with retroviruses,
 XX cultured and implanted under the renal capsule of mice.
 XX Reconstructions were harvested 5 weeks later, when they showed
 CC signs of distress from the tumour burden. Metastatised tumours were
 CC isolated from a site outside the renal capsule. RNA was isolated
 CC from primary tumours and metastases, reverse transcribed and
 CC subjected to differential display PCR. The present sequence was analysed
 CC to obtain metastatic sequences (e.g. the present sequence). The
 CC related sequences were used to screen a biological sample for
 CC metastasis, or treat malignant or non-malignant disorders, e.g.
 CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can
 CC be used to detect, diagnose and treat disorders related
 CC to metastasis, or treat malignant or non-malignant disorders, e.g.
 CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can


```

RESULT 11
AA096055
XX ID AA096055 standard; cDNA to mRNA; 2115 BP.
XX XC
XX DT 22-JAN-1996 (first entry)
XX DE Sequence encoding melanoma associated antigen gp100.
XX KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
XX KW Identification; tumour; gp100; ds.
XX OS Homo sapiens.
XX FH
XX FH Key Location/Qualifiers
XX CDS 22..2007
XX FT /tag= a
XX FT /product= Melanoma associated antigen gp100.
XX FT misc_signal 181
XX FT /tag= b
XX FT /tag= c
XX FT /label= Transmembrane domain.
XX FT misc_feature 1792..1870
XX FT /tag= d
XX FT /misc_feature 262..264
XX FT /misc_binding 337..339
XX FT /misc_binding 352..354
XX FT /tag= e
XX FT misc_binding 982..984
XX FT /tag= f
XX FT misc_binding 1725..1725
XX FT /tag= h
XX PN EP668350-A1.
XX PD
XX PD 23-AUG-1995.
XX PR 14-FEB-1995; 95EP-0200348.
XX PR 21-DEC-1994; 94EP-0203709.
XX PR 16-FEB-1994; 94EP-0200337.
XX PA (AUKU ) AKZO NOBEL NV.
XX XX
XX XX Adema GJ, Fijdor CG;
XX XX WPI: 1995-284790/28.
XX DR P-PSDB; AAR78646.
XX XX
XX XX Melanoma associated antigen gp100 - used in vaccines and for the
XX XX detection of tumours
XX XX
XX XX Claim 2; Page 19-22; 40pp; English.
XX CC
XX CC Immunogenic peptides derived from the melanoma associated antigen
XX CC may be used in the production of vaccines. Nucleotide sequences
XX CC encoding the immunogenic peptides may be used as primers and probes
XX CC in the detection of melanoma cells. Tumour infiltrating lymphocytes
XX CC capable of binding to the melanoma associated antigen can be
XX CC cultured ex vivo and returned to melanoma particles, and when
XX CC radiolabelled, they may be used to identify tumour deposits.
XX XX
XX XX Sequence 2115 BP; 469 A; 587 C; 575 G; 484 T; 0 other;
XX XX
XX XX Alignment Scores:
XX XX Pred. No.: 2,16e-38
XX XX Score: 627.50
XX XX Percent Similarity: 96
XX XX Best Local Similarity: 27.08%

```

```

Query Watch: 20.50% Indels: 173
DB: 16 Gaps: 19
US-09-943-075a-2 (1-572) x AA096055 (1-2115)
Oy 7 ValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGlnInAlaIaIys-----Arg 24
Oy 46 CTTCTTCATTGGCTGTGTATAGGTGCTTCTCTGGCTGTGGGCTTCAAAAGTACCCAGA 105
Oy 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnArgInLeu 44
Db 106 AACCAAGGACTGGCTTGGT-----GTCTCANGGCAACTC 138
Oy 45 ArgGlyTyrSerSerAspGluAsnGlnTyrPheGluGlnLeuTyrProValTyrArgArg 64
Db 139 AGA-----ACCAAGCCTGGAACAGGCAGCTGTATCCAGAGTGACAA-- 180
Oy 65 GlyGluGlyArgTyrLysAspSerTrpGluGlyGlyArgValGlnAlaAlaLeuThrSer 84
Db 181 ----GAGCCCAAGAGCTGTCTGGAGAGGTGTCTGAGTGTCCCTCANGTACGTATG 237
Oy 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
Db 238 GATGGCTACACTGATGTCGCAATGGCTCTCTCTATTGCTTGAACCTCCCTGGA 297
Oy 105 CysGlnGlyGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
Db 298 AGCCAAAGAGTATCCCATGGGCGAGGTATCTGGGTCCACCAATCCATCATCGATGGG 357
Oy 124 LeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrThrGlyValAspAspGluAsp 143
Db 358 AGCCAGGTGTGGGAGGAGCAGCAGTGTATCCCGAGGAAACT-----GAGCATGCC 408
Oy 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163
Db 409 -----TGCATCTCCCTGATGATGATGGTGGACCTTGGCTGGACCTTGGCC 438
Oy 164 ArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeuGlyGlnTyr 183
Db 439 TCTGGCTCTGGTCTCAGAGAGAGAGCTTTGTTATGTCTGGAGAGCTGGGGCCAAATC 498
Oy 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
Db 499 TGGCAAGTTCTAGGGGCCAGTGTCTGGCTGAGCATTTGGAGAGCGAGGCAATGCTG 558
Oy 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222
Db 559 GGCACACACACCATGGAGTACTGTCTACCATCGCGGGGATGCCGGAGCTATGGCT 618
Oy 223 IleSerLysValLysAspValTyrValIleThrAspGlnGlnLeuProIlePheValThr 242
Db 619 CTTGCTCATCCAGCTCAGCCTTACACATTACTGACCAAGTGGCTTTCTCGTGGAGCTG 678
Oy 243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
Db 679 TCCCACTTGGCGGCTTGGATGGAGGAGCAGCAGCTCTCTGGAATATGCCCTCTGACC 738
Oy 263 PheAspValLeuIleHisAspProSerThrIlePheLeuAsnTyrSerAlaIleSerTyrIys 282
Db 739 TTTGCTCTCCAGCTCATGCCCGCAGTGTCTGCTGAGCTGAAGCTGACCTCTCTACACC 798
Oy 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
Db 799 TGGGACTTTGGAGAGTATGGAGACCTCATCTCTCGGGGCACTTTGTGTGCTCATCATCT 858
Oy 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320
Db 859 TACCTGGAGCTGGCGGAGTCACTCTGCTGCTGAGCTGGCAGCTGCCATCTCTCTCACC 918
Oy 321 -----GlyProCysPro----- 324
Db 919 TCTCTGGCTCTCTCCCGAGTTCCAGGGCAACACAGATGGGCAAGGCCAACTGAGAGGCC 978

```

QY 324 ----- 324
 Db 979 CTTACACACAGCTGGCCAGTGGCTCTACTACAGAGTGTGGGTACTACACCTGGCTCAG 1038
 QY 325 SerProThrProSerSerThr----- 334
 Db 1039 CGGCCAATCGAGAGCCCTCTGGAACACACATCTGTGAGGTGCCAACCTGAAGTCATA 1098
 QY 335 SerProSerProIaSerProSerProThrProThrLeuSerProSerProSerLeuMet 354
 Db 1099 ACGCTGACCTGTGCAGATGCGACCTCGACAGACACAGGTATGACACCTCGAGAGGTG 1158
 QY 355 PTO-----ThrGly 357
 Db 1159 CCAGTTTCAGAGTCTCATGGGTACCACTGCGCAGAGATGTCACACTCCAGAGCTACAGGT 1218
 QY 358 TyrLySerMetLeuSerAspIle----- 366
 Db 1219 ATGACACCTGCAAGGTATCAATGTGGCTCTTCTGGAACACAGCTGCACAGGTACCA 1278
 QY 366 ----- 366
 Db 1279 ACTACAGATGGTGGAGACACAGCTAGAGAGCTACCTATCCCTGAGGCTGGAAGGTCCA 1338
 QY 367 -----SerAnglu----- 369
 Db 1339 GATGCCAGCTCAATCATGTCTACGGAAGATTACAGAGTTCCCTGGGCCCTCGCTGGAT 1398
 QY 370 -----AsnCysArgIleAsnArg 375
 Db 1399 GGTACAGCCACTTAAAGCTGGTGAAGACAGAGTCCCTGGATGTGTCTGTATCGA 1458
 QY 376 TTTGlyTyrPheArgAlaThrThrIleValAspGlyIleLeuGluValAsnIleIle 395
 Db 1459 TATGGTTCCTTTCCGTCACTGACCTGGACATGTCCAGGTAITGAAAGTCCGAGATCTCG 1518
 QY 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
 Db 1519 CAGGCT-----GTGCGTCCGGTGAAGGGGAT-----GCATTTGAGCTGACT 1560
 QY 416 ValThrCysLyseIleValaThrProThrIleValaThrIleLeuSerAspProThrCys 435
 Db 1561 GTGTCTCCACAGCGGGGTGCCCGGAGCATCTGCATGGAGATCATCTGCCAGGCTGC 1620
 QY 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValaIleAspGlyLeuSerLeuSer 455
 Db 1621 CAGCCCTCGCCACCGGTGTGGTCCAGCTGTGTCTACCGACCCAGCTGCCAGCTGGT 1680
 QY 456 ValArgArgAlaPhe--AsnGlySerCylThrCysValAsnPheThrIleGlyAsp 474
 Db 1681 GTGCACCAAGTACTCAAGGTTGGCTCGGGAGCATCTGCCTCAATGTCTCTGGCTGAT 1740
 QY 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspGly--- 493
 Db 1741 ACCAACAGCTGGAGTGGTGTGCGACGCCACCTATATGCTGTCTCGAGACGCGCTT 1800
 QY 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
 Db 1801 GGGCAGGTTCGCTGATCGGTGGCATCTGTG-----CTGGGTGTGATGGCT 1845
 QY 514 ThrMetValThrLeuLeuTyrLysLys 523
 Db 1846 GTGTCTTGTGCACTCTGATATAGCGGC 1875
 RESULT 12
 ID AAH43500 standard; cDNA; 2130 BP.
 XX AAH43500
 XX AAH43500;
 XX 13-DEC-2001 (first entry)
 XX Human melanoma antigen gp100 coding sequence.

XX Major histocompatibility complex; MHC; human; melanoma antigen; gp100;
 KW HLA-A2 binding domain; mutation; antigen presenting cell; vaccine;
 XX immune effector cell; cancer; antibody; ss.
 OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 22..2007
 TT /*tag= a
 PN /product= "gp100"
 PD WQ200170767-A2.
 PD 27-SEP-2001.
 XX 19-MAR-2001; 2001WO-US08919.
 XX 20-MAR-2000; 2000US-190750P.
 PR 12-DEC-2000; 2000US-255019P.
 XX (GENZ) GENZYME CORP.
 PA Nicolette CA.
 PI WPI: 2001-611469/70.
 DR P-PSDB; AAH47500.
 XX Novel synthetic compounds useful for stimulating an immune response in
 PT a subject and as components of anti-cancer vaccines, are designed to
 PT enhance binding to major histocompatibility complex molecules -
 XX Disclosure: Page 60-63; 67pp: English.
 CC This sequence encodes human melanoma antigen gp100. Peptides of the
 CC invention based on the sequence of residues 209-217 of human melanoma
 CC antigen gp100, which represents the putative HLA-A2 binding domain,
 CC are designed to enhance binding to major histocompatibility complex
 CC (MHC) molecules and to enhance immunoregulatory properties relative to
 CC confer tighter binding to the MHC. These peptides are useful for
 CC inducing an immune response in a subject. Where they are delivered in
 CC the context of an MHC molecule which presents the compound on the
 CC surface of an antigen presenting cell. The peptide sequences are useful
 CC as components of anti-cancer vaccines and to expand immune effector cells
 CC that are specific for cancers characterized by expression of the
 CC melanoma antigen gp100. The compounds are useful for treatment of
 CC diseases such as cancer, in particular against human melanoma and for
 CC generating antibodies that specifically recognize and bind the compounds.
 XX Sequence 2130 BP; 484 A; 587 C; 575 G; 484 T; 0 other;
 Alignment Scores:
 Length: 2,186-38
 No.: 2130
 Score: 627.50
 Percent Similarity: 41.85%
 Best Local Similarity: 27.08%
 Mismatches: 205
 Indels: 173
 Gaps: 19
 Query Match: 22
 DB: 20.50
 US-09-943-075a-2 (1-572) x AAH43500 (1-2130)
 QY 7 ValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
 Db 46 CTTCTTCATTGGCTGTGATAGGTGCTTTGCTGGCTGTGGGGGCTACAAAGATACCAGA 105
 QY 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGluAsnGlnLeu 44
 Db 106 AACGAGACCTGGCTTGGT-----GTCTCAAGGACATC 138
 QY 45 ArgGlyTyrPheSerAspGluAsnGlnTyrPaspGluGlnLeuTyrProValTyrpArg 64
 Db 139 AGA-----ACCAAGCCTGGAAACAGGAGCTGTATCTCCAGATGGACA--- 180

[illegible]

OY 216 HisGly---ArgAlaTyrIleProIleSerIysValIysAspValTyrValIleThrAsp 234
 DB 612 CGGGATCCGGAGCTATGCGCTTCTGCTTATTCCTCCAGCTCAGCTTCACCATCTAGAC 671
 OY 225 GluIleProIlePheValThrMetTyrGlnIysAspAspArgAsnSerSerAspGluThr 254
 DB 672 CAGTGGCTTTCTCGCTGGAGCGCTGCCAGTTCCGGCGCTTGATGGAGGAACAAGCAG 731
 OY 235 PheLeuArgAspLeuProIlePhePheAspValLeuIleHisAspProSerHisPheLeu 274
 DB 722 FTCTGGAGAAATCAGCTCTGACCTTTCGCTGCTCCAGCTCATGAGCCAGTGGCTATCTG 791
 OY 275 AsnIleSerAlaIleSerTyrIleTyrPheIysAspAsnThrGlyLeuPheValSer 294
 DB 792 GCTGAGCTGCTCTCTACACCTGGAGCTTTGGAGACAGTAGTGGAAACCTCATCTCT 851
 OY 295 AsnAsnIleThrLeuAsnHisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThr 314
 DB 852 CGGGCACTTGTGTGCTACTCATATTACTTGAGGCTCGCCAGCTCACTGCGCCAGGTGGTC 911
 OY 315 ValGlnThrAlaValPro---GlyProCysPro--- 324
 DB 912 CTGACAGCTGCGCATCTCTCTCACTCTGCTGTGGCTCTCCCGCAGTTCAGGCAACACAGAT 971
 OY 324 ----- 324
 DB 972 GGGCACAGGCCACTCGAGAGGCCCTTAACACACAGCTGGCCCAAGTGCCTACTACAGAA 1031
 OY 325 -----SerProThrProSerProSerSerSerThr----- 334
 DB 1032 GTTGTGGGTACTACACTGTGTCNGGGCGCAACTGCGAGAGCGCTCTGGAAACCAATCTGTG 1091
 OY 335 -----SerProSerProAlaSerSerProSerProThrLeu 346
 DB 1092 CAGTGTCCAACTCAAGTCAATAAGCACTGCACCTGTGCAGATGCCAACTGCAGAGAGC 1151
 OY 347 SerThrProSerProSerLeuMetPro----- 355
 DB 1152 ACGGTATGACACTCGAGAAGGTGCCAGTTTCAGAGTCACTGGTACACACTGCCAGAG 1211
 OY 356 -----ThrGlyTyrIleSerMetGluLeuSerAspIle----- 366
 DB 1212 ATGTCAACTCCAGAGGCTACAGTATGACACCTCGCAGAGGATATCAATTGTGCTGCTTCT 1271
 OY 366 ----- 366
 DB 1272 GGACCACTGCTCAGAGTACACACTACAGAGTGGGTGGAGCAACAGCTAGAGAGCTA 1331
 OY 367 -----SerAsnGlu----- 369
 DB 1332 CTTATTCCTGAGCTGAAGTCCAGATGCCAGTCAATCATGTCTACGAAAGATTATCA 1391
 OY 369 ----- 369
 DB 1392 GGTTCCTCGGCGCCCTGCTGGATGTACAGCCACTTAAAGCTGGTGGAGAGCAAGTCT 1451
 OY 370 -----AsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAsp 387
 DB 1452 CCCTCGGATTGTGCTGATGCAATATGTTCTTTTCGGTCACTCGACCTGGACCTGTCCAG 1511
 OY 388 GlyIleLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro 407
 DB 1512 GGTATTCGAAAGTCCGAGATCTCTCAGCT-----GTGCGCTCCGGTGGAGGG 1559
 OY 408 AspAsnSerLeuMetAspPheIleValThrCysLysGlyValThrProThrGluAlaCys 427
 DB 1560 GAT-----GATTTTGTAGCTGATGTCTCTCGCAAGGGGGGCTGCCAAGGAGAGCTGCG 1613
 OY 428 ThrIleIleSerAspProThrCysIleIleIleGlnAsnArgValIcysSerProValAla 447
 DB 1614 ATGGAGATCTCTCGCAGAGGTGACAGCCCTGCGCCAGCGCTGTGACAGAGCTGTGCTA 1673
 OY 448 ValAspGluCysLeuLeuSerLeuValArgAlaPhe---AsnGlySerGlyThrTyr 466

DB 1674 CCCAGCCACCTCCAGCTGAGCTGTTCTGACCAAGATCTGAGAGGTGGCTCGGGACATAC 1733
 OY 467 CysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSer 486
 DB 1734 TGCTCAATGTCTCTGCTGATACCAACAGCCTGGCAGTGGTCAGCACCCAGCTTATC 1793
 OY 487 IleProGlyIysAspLeuGly---SerProLeuArgThrValAsnGlyValLeuIleSer 505
 DB 1794 ATGCGCTGTGTCAGAACAGCCCTGGCAGAGTCCCTGATCTGGCAGCTCTG----- 1847
 OY 506 IleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrIleLys 523
 DB 1848 -----CTGTGTTGATGGCTGTGCTCTGCTCATCTCTGATATATAGCGC 1892
 RESULT 15
 T: AAS45525 standard; cDNA; 2172 BP.
 AC AAS45525;
 XX 18-DEC-2001 (first entry)
 DT DNA encoding Melanoma antigen CDNA25.
 XX Human: MARR-1; immunogenic; melanoma antigen recognised by T lymphocyte;
 KW diagnostic; therapeutic; vaccine; melanoma; in vivo tumour recognition;
 KW in vivo tumour rejection; ss.
 XX Homo sapiens.
 XX US527/0778-B1.
 XX 07-AUG-2001.
 PD 12-MAR-1999; 99US-0267439.
 PF 05-MAY-1998; 9805-0073138.
 PK 22-APR-1994; 9405-0231565.
 PK 05-APR-1995; 9505-0417174.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Kawakami Y, Rosenberg SA;
 PI WPI; 2001-595403/67.
 DR P-PSDB; A028912.
 XX Immunogenic peptide useful in vaccines comprises specific amino acids
 PT of new melanoma antigen recognised by T lymphocytes -
 XX Example 3; Figure 4; 73pp; English.
 XX The invention relates to a novel immunogenic peptide comprising 5-20
 CC amino acids and a carrier, and a method of using the peptide in the treatment
 CC of melanoma (MARR-1). The peptide sequence contains at least one amino
 CC acid modification of MARR-1. The peptide is used in diagnostic and
 CC therapeutic methods as an immunogen or vaccine to prevent or treat
 CC melanoma, and for in vivo tumour recognition and rejection. AAS45524-
 CC AAS45528 represent MARR-1 coding sequences, PCR primers, and related
 CC sequences of the invention.
 XX Sequence 2172 BP; 512 A; 594 C; 578 G; 488 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2,44e-38 Length: 2172
 Score: 627.00 Matches: 174
 Percent Similarity: 41.03% Conservative: 96
 Best Local Similarity: 16.44% Mismatches: 126
 Every Match: 2.46% Gaps: 182
 DB: 19
 US-09-943-075A-2 (1-572) x AAS45525 (1-2172)

```

QY 17 LeuProLeuInAlaAlaLysArgPheAlaAspVal----- 28
DB 12 TTACCAATCGCGACCGGAAGAACAACATGATCTGGTCTAAAGAGATGCTTCTTCAT 71
QY 29 -----LeuGlyHisGluInTyPro----- 36
DB 72 TTGGCTGTAGTAGTGCTTTCTGGCTGTGGGGCTACAAAGATGCCCAAAACAGAC 131
QY 37 HisMetArgGluAsnAsnGlnLeuValGlyTyrPheSerAspGluAsnGluTyrAspGlu 151
DB 132 TGGCTGTGGTGTCTCAAGGCAACTGAGA-----ACCAAGGCTGGAGACAG 176
QY 57 GlnLeuTyProValTyrPheArgGlyGluGlyAlaGlyTyrPheLysAspSerTyrGluGly 76
DB 177 CACCTGTATCCAGAGTGACA-----GAAGCCAGAGACTTGACTCTGGAGAGTGGT 230
QY 77 ArgValGlnAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96
DB 231 CAAGTGTCCCTCAAGGTCAGATGATGGAGCCAGGTGTGGCTACACTGATGTGTGCAAAAGCTCTCTTC 290
QY 97 ValValAsnLeuValPhePheArgGlyGlyGluGlyAspAlaAsnGlyAsnIleValTyr 216
DB 291 TCTATTGCCCTCAACTCTCCCTGGAGGCCAAAAGCTATTGCCAGATGGCAGGTATCTGG 350
QY 117 GluArgAsn---CysSerSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrp 135
DB 351 GTCACATACCATCATCATATGGAGCCAGGTGTGGGAGCCAGCCAGTGTATCCCGAG 410
QY 136 ThrThrGlyAlaAspAspGluAspTrpGluAsnThrSerGlnGlyGlnHisLeuArg 155
DB 411 GNAACCT-----GACCATGCC-----TGCAATC 431
QY 156 PheProAspGlyLysProPheProArgProHisGlyArgLysLysTrpAsnPheValTyr 175
DB 432 TTCCGTATGGTGCACCTGCCCATCTGGCTCTGGTCTCAGAGAGAAGCTTTGTTAT 491
QY 176 ValPheHisThrLeuGlyGlnTyPheGlnIleLysLeuGlyGlnCysSerAlaArgValSer 195
DB 492 GTCTGGAAGCTTGGGGCCAATCACTGCAATTTCTAGGGGGCCAGTGTCTGGCTGAGC 551
QY 196 IleAsnThrValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgTrp 215
DB 552 ATTTGGGACGAGGCAAGTGTCTGGGCACACACACACACATGGAAGTGACTGTCTACCATCGC 611
QY 216 HisGly---ArgAlaTyrIleProIleSerIleValLysAspValTyrValIleThrAsp 234
DB 612 CGGGATTCGGGAGCTATGGCTGTCTGCTTCCTACCTCAGCTCAGCTCTCCATCTACGAC 671
QY 235 GlnIleProIlePheTyrMetGlyTrpGlnTyAsnAspArgAsnSerSerAspIleThr 254
DB 672 CAGGTGGCTTTCTCCGTGACCGGTGCTCCAGATTGGCGGGCTTGGATGGAGGACACAGAC 731
QY 255 PheLeuArgAspLeuProIlePhePheAspValIleuIleAspProSerHisPheLeu 274
DB 732 TTCTCGAAGATCAGCTCTGACCTTTGGCTTCAGCTCCAGCTCATCAACCCAGTGCATCTG 791
QY 275 AsnArgAlaIleSerTyrIleTyrTrpAsnPheGlyAspAsnThrGlyLeuPheValSer 294
DB 792 GCTGAAGCTGACCTCTCCCTACCTGGGAGTTTGGAGACAGTGTGGAACCCCTGATCTCT 851
QY 295 AsnAsnHisThrLeuAsnHisThrValLeuAsnGlyThrPheAsnPheAsnLeuThr 314
DB 852 CGGGCACTTGTGGTCACTCATATTACTTGGAGGCTGGCCAGCTCACTGGCCAGTGGTC 911
QY 315 ValGlnAlaValPro-----GlyProCysPro----- 324
DB 912 CTCAGGCTGCCATCTCTCTACCTTCCTGGCTCTCCCTCCAGTTCAGGACACACAGAT 971
QY 324 ----- 324
DB 972 GGGCACAGGCCAACTGCAGAGGCCCTTAACACACAGCTGGCCAAAGTGCCTACTACAGAA 1031

```

Search completed: April 8, 2003, 14:52:59
Job time : 381 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 9, 2003, 02:18:55 ; Search time 59 Seconds

Title: US-09-943-075A-2
Perfect score: 3061
Sequence: 1 MESLGGVLLIAGLPLQ.....PFSRGREKDPPLQKPKML 572

Scoring table: BLOSUM62
Gapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters: -p2n model
-model frame_plus_p2n model
-db-issued Patents NA -UCDB-175 -runset_07042003_155631_17949 -app_query.fasta_1.711
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi
-LIST=45 -LOCAL=200 -THR_SCORE=pept -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09943075A -ECCG=1 -40 -runset_07042003_155631_17949 -ICPU=3
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YFUTOP=6 -YFUTOP=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMP.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMP.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMP.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMP.seq:*
5: /cgn2_6/ptodata/1/ina/5E_COMP.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	DB ID	Description
1	2700	88.2	2213	4	US-09-383-586-27
2	2163	70.7	2669	1	US-08-594-031-91
3	2163	70.7	2669	1	US-08-594-031-91
4	941	30.7	620	1	US-08-594-031-113
5	729	23.8	524	1	US-08-594-031-105
6	625	20.5	473	1	US-08-594-031-115
7	625	20.5	2130	4	US-08-594-031-115
8	627	20.5	2112	2	US-08-594-031-115
9	627	20.5	2112	2	US-08-594-031-115
10	627	20.5	2112	2	US-08-594-031-115
11	627	20.5	2112	2	US-08-594-031-115
12	524	17.1	335	1	US-08-594-031-94

13	524	17.1	335	1	US-08-594-031-94	Sequence 94, Appl	
14	524	17.1	335	1	US-08-594-031-94	Sequence 96, Appl	
15	524	17.1	335	1	US-08-594-031-123	Sequence 123, Appl	
16	504	16.5	354	1	US-08-594-031-114	Sequence 114, Appl	
c	17	41	15.4	286	1	US-08-594-031-121	Sequence 121, Appl
18	417	13.9	262	1	US-08-594-031-93	Sequence 93, Appl	
19	418	13.7	309	1	US-08-594-031-106	Sequence 106, Appl	
c	20	419	13.4	260	1	US-08-594-031-119	Sequence 119, Appl
21	368	12.7	270	1	US-08-594-031-107	Sequence 107, Appl	
22	368	12.7	270	1	US-08-594-031-109	Sequence 109, Appl	
c	23	364.5	11.9	335	1	US-08-594-031-111	Sequence 111, Appl
24	356	11.6	223	1	US-08-594-031-116	Sequence 116, Appl	
25	350	11.4	240	1	US-08-594-031-118	Sequence 118, Appl	
c	26	317.5	10.4	217	1	US-08-594-031-112	Sequence 112, Appl
27	294	9.6	190	1	US-08-594-031-95	Sequence 95, Appl	
c	28	293	9.6	231	1	US-08-594-031-130	Sequence 120, Appl
29	279	9.1	247	1	US-08-594-031-117	Sequence 117, Appl	
c	30	216	7.1	224	1	US-08-594-031-122	Sequence 122, Appl
31	135	4.4	18596	4	US-09-318-448-11	Sequence 11, Appl	
c	32	131.5	4.3	3004	1	US-08-276-213-6	Sequence 6, Appl
33	130	4.2	1293	1	US-08-604-9138-12	Sequence 12, Appl	
34	129	4.2	1293	4	US-09-032-469-5	Sequence 5, Appl	
35	129	4.2	14148	4	US-09-032-469-7	Sequence 7, Appl	
c	36	128.5	4.2	377	2	US-08-332-766A-1	Sequence 1, Appl
37	126.5	4.1	3337	1	US-08-072-610-1	Sequence 1, Appl	
c	38	126.5	4.1	3337	2	US-08-719-822B-1	Sequence 1, Appl
39	126.5	4.1	3337	4	US-09-032-458-1	Sequence 1, Appl	
c	40	125	4.1	2500	1	US-08-034-650-9	Sequence 9, Appl
41	123	4.1	18596	4	US-08-318-448-11	Sequence 9, Appl	
42	123	4.1	14060	3	US-08-658-136-4	Sequence 4, Appl	
43	124	4.1	12912	2	US-08-460-751-1	Sequence 1, Appl	
44	123	4.0	11707	4	US-09-136-574A-1	Sequence 1, Appl	
45	120	3.9	4800	3	US-09-106-638-1	Sequence 1, Appl	

ALIGNMENTS

RESULT 1
US-09-383-586-27
; Sequence 27, Application US/09383586
; US-09-383-586-27
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Kumble, Anand
; APPLICANT: Kumble, Anand
; TITLE OF INVENTION: Cytokine gene
; TITLE OF INVENTION: Cytokine gene
; TITLE OF INVENTION: Cytokine gene
; FILE REFERENCE: 11000.1037C1
; CURRENT APPLICATION NUMBER: US/09/383.586
; NUMBER OF SEQ ID NOS: 38
; CURRENT FILING DATE: 1999-08-26
; SEQ ID NOS: 1-38
; SEQ ID NOS: 1-38
; LENGTH: 2213
; TYPE: DNA
; ORGANISM: Mouse
US-09-383-586-27

Alignment Scores:
Pred. No.: 7.91e-257
Score: 2700.00
Length: 2213
Matches: 507
Conservative: 20
Mismatch: 42
Best Local Similarity: 88.79%
Query Match: 88.21%
Gaps: 1
DB: 4

US-09-943-075A-2 (1-572) x US-09-383-586-27 (1-2213)

QY 1 MetcLuserLeuCyGclYvaLeuVaIphLeuLeuAlaAlaGlyLeuProLeuGln 20
|||||

Db 23 ATCGAAAGTCTCTCGGGGCTCGGGGATTTCTGCTGGCTGACGAGCTGCTCCAG 82
 Qy 21 AlAlaIyAsrAheArAspVaIlleuGlyHisGluGlnTyrProAspHisMeArgGlu 40
 Db 83 CTGCGCAAGCGGATTTCTGATGATGCTGGCGCATGAACAGTATCCCAATACATGAGAG 142
 Qy 41 AsnAsnGlnLeuAspGlyTyrProSerAspGlnAspGlnTyrProAspGlnLeuTyrPro 60
 Db 143 CACAACCAATTAAGTCTGGTCTTTCGGATCAAAATGAATGGATGAACACCTGTATCCA 202
 Qy 61 ValTyrArgArgGlyGluGlyArgTyrPlyAspSerTrpGluGlyArgValGlnAla 80
 Db 303 GTGTGGAGGGAGGAGCGCGAGTGGAGAGCATCTGGAGAGAGGCGCTGTGGAGCCA 262
 Qy 81 AlaLeuThrTrpAspSerProAlaLeuValGlyCysAsnIleThrValValAsnLeu 100
 Db 263 GTCTGACACAGTACATCAAGCGCTCTGCTGGGTTCCTCAATATCACCTTTGTGTGAACCTG 322
 Qy 101 ValPheProArgCysGlnIyAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
 Db 323 GTGTTCCTCCAGATCCCAAGAGAGATCTCTATAGCAATATCTCTATGAGAGAGCTCC 382
 Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140
 Db 383 AGGAATGATTGGAGCTGACCTCTGACCTGACCTGATGCTACACCTGGAGCGGCGAGAT 442
 Qy 141 AspGluSerTrpGlnAsnAsnThrSerGlnGlyGlnHisLeuAspPheProAspGlyLys 160
 Db 443 CATGGTGAAGTGGCAAGTGGCACCACCGAGCGAGCATCTCAGTTCCTCCGACGAGGG 502
 Qy 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180
 Db 503 CCTCTCTCTGCGCCCTCGTGGAGAGAGTGGCTGTCTACCTCTTTCACACACTT 562
 Qy 181 GlnGlnTyrPheGlnIyAspLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
 Db 563 GGCCAGATATTTCCAAAACATGGGTGCTGTTCAGCAGCGGGTTCTATAAACACAGTCAAC 622
 Qy 201 LeuThrValGlyProGlnValAspGluValIleValPheThrArgHisGlyArgAlaTyr 220
 Db 623 TTGACATGCGCCCTCAGGTATGGAAGTGAAGTGAAGTCTTTTGGAGATACGCGCGGCATAC 240
 Qy 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
 Db 683 ATTCACATCTCGAAGGTGAAGTGTGTATGTGTACAGATCAGATCTGTATTCGTG 742
 Qy 241 ThrMetTyrGlnIyAspAspArgAsnSerSerAspGlnThrPheLeuArgAspLeuPro 260
 Db 743 ACCATGTGCCAGAAGATGACAGAACTGTGCTGATGAGATCTTCTCCAGAGACTCCCC 802
 Qy 261 IlePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
 Db 803 ATGCTCTCTCATGATCTCATGATGATGATCCAGCACTTCTCCACACTTCCCATTC 862
 Qy 281 TyrLysTrpAsnPheGlnIyAspAsnThrGlyLeuPheValSerAspAsnHisThrLeuAsn 300
 Db 863 TACAAAGTGGAACTTTGGGGACAACTGGCTGTGTGTTCTCCCAACAATCAACATTTGAAT 922
 Qy 301 HisThrValLeuAsnGlnIyPhePheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 Db 923 CACACTATGCTCATGGAACCTTCAACCTTCACTCACCCTGCAACTGCAGTGCCTC 982
 Qy 321 GlyProCysProSerProThrProSer-----ProSerSerSerThrProSerPro 338
 Db 983 GGCCATGCGCTCCCTCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
 Qy 339 AlaSerSerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 358
 Db 1043 CCAGCCCTCACTCTGGCCACCATATTCACACACTAGCCCTCTTTTAATGCTCTAGTGTAC 1102
 Qy 359 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysAtyGlnIleAsnArgTyrGlyTyr 378
 Db 1103 AATATCATGAGGCTGATGTGATTTCTCATGTATGATGATGATGATGATGATGATGATGAT 1162

Qy 379 PheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAla 398
 Db 1163 TTTCAGAGCCACCATCAATTTGTAGAGGGGATCTCGAAGTCAAGCATCATGACAGATAGCA 1222
 Qy 399 AspValProIleThrIleThrLeuGlnProAspAsnSerIleuMetAspPheIleValThrCys 418
 Db 1282 GATGTCCTCATGCGCATACACAGCTGCGCACTGCTGATGATGATGATGATGATGATGATGAT 1282
 Qy 419 LysGlyAlaThrProThrIleGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 438
 Db 1283 AAAGGGGCCACCCCATTTGGAAGCTGTATCATCATCTCCGACCCACCTGCGAGATGCGC 1342
 Qy 439 GlPheArgValCysSerProValAlaValAspGluValLeuLeuLeuValArgSer 458
 Db 1343 CAGAACCGGGTCTCGACGCTGTGGATGGCTGTGCTGTGCTGTGCTGTGCTGTGAGAGA 1402
 Qy 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478
 Db 1403 GCTTCATATGGGTGTGGCACTACCTACTGTGTGAATTTCTCTCTGTGAGATGATGAGCTG 1462
 Qy 479 AlaLeuThrSerAlaLeuIleSerIleThrProGlyLysAspLeuGlySerProLeuArgThr 498
 Db 1463 GCGCTCAGCAGACCCCTGATCTATCCCTGCGCAAGACCCAGACTCCCTCTGAGAGCA 1522
 Qy 499 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIle 518
 Db 1523 GTGAATGGGTGTGATCTCTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1582
 Qy 519 LeuLeuTyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValVal 538
 Db 1583 TTGCTGTACAAAACACAAAGCGTACAGCCCAATAGAACTGCGCCAGGAACAGGTGTC 1642
 Qy 539 LysGlyLysGlyLysValPheLeuSerHisAlaLysAlaLysProPheSerArgGlyAsp 558
 Db 1643 AAGGSCAAAGGCGTGTGCTCTCTCAGCCACCGCAAGCCGCCCTTCTTCGAGAGAGAC 1702
 Qy 559 ArgGlyLysAspProLeuLeuGlnAspLysPro 569
 Db 1703 CAGAGAGAGTCCNTTCTCTCAGGACAGCA 1735
 RESULT 2
 US-08-594-031-91
 ; Sequence 91, Application US/08594031
 ; Patent No. 5783182
 ; Grant No. 5783182
 ; APPLICANT: THOMPSON, Timothy C.
 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 ; NUMBER OF SEQUENCES: 175
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BAKER & BORTS, L.L.P.
 ; STREET: 1299 Pennsylvania Avenue, N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20004-2400
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; SOFTWARE: SYBASE
 ; SOFTWARE: FASTSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/594, 031
 ; FILING DATE: 30-JAN-1996
 ; CLASSIFICATION: 435
 ; PUBLICATION NUMBER: 60/006,838
 ; APPLICATION NUMBER: 60/006,838
 ; FILING DATE: 16-NOV-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Remenick, James
 ; REGISTRATION NUMBER: 36,902
 ; REFERENCE/DOCKET NUMBER: 0A146-0110
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-639-7700
TELEFAX: 202-639-7850
TELEX:
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 3669 base pairs
TYPE: nucleic acid
FEATURES: repeat region
TOPOLOGY: linear
MOLECULAR TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
SOURCE:
US-08-594-031-91

Alignment Scores:
Pred. No.: 1,246-203 Length: 2669
Score: 2163.00 Matches: 394
Percent Similarity: 82.57% Conservative: 85
Query Similarity: 82.57% Conserved: 85
Query Match: 70.66% Indels: 14
DB: 1 Gaps: 1

US-09-943-075a-2 (1-572) * US-08-594-031-91 (1-2669)

```

QY 1 MetGluSerAspGlyValLeuValPheLeuLeuAlaLaGlyLeuProLeuGln 20
DB 1 ATGGAATGCTCTACTATTTCCTGGGATTTCTCTGCTGCTGCAGATTCCTCATGTGAT 151
QY 21 AlaAlaValArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
DB 152 CCGCCCAACAGATTTCATGATGCTGGGCAATGAAGACCTTCCTCTACATGAGGAG 211
QY 41 AsnAspGluLeuArgGlyTyrSerSerAspGluValPheArgGluGlnTyrPro 60
DB 212 CACATCAATTAATGCTGCTCTCTGATGAAATGACTGCAATCAAAATCTCTACCA 271
QY 61 ValTyrArgGlyGlyGluArgTyrPheAspSerTrpGluGlyValGlnAla 80
DB 272 GTGTGGAGGGGAGCATGAGGAGGAATCTCTGGAGGAGGCGCTGTCCAGGCG 331
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnLeuThrPheValValAsnLeu 100
DB 332 GTCTGCACAGTGAATCACCAGCCCTCGTGGGCTCAATATACATTTGGCGGTGAACCTG 391
QY 101 ValPheProArgCysGlnIysGluAspAlaAsnGlyAsnIleValTyrGluArgAspCys 120
DB 392 ATATTCCTGATGATGCAAGAGATTCGCAATGCGACATAGCTATGAGAGAGATCGC 451
QY 121 ArgSerLeuLeuLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140
DB 452 AGAATAGAGCTGGTATTCTCTGCTGCTCATATGTTTACACATGGACAGCATGGTCCAGAG 511
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnIleLeuArgPheProAspGlyLys 160
DB 512 GACAGTGAAGGGGAAATGCGACCGGCCAACCATCACTACGCTTCCTCTGCTGGGAA 571
QY 161 ProPheProArgProHisGlyArgTyrAspAsnValTyrValPheHisThrLeu 180
DB 572 CTTTTTCTTACACACCCCGGAGGAGCATGGAATTTTCATCTACGCTCTCCACACACT 631
QY 181 GlyClnTyrPheGlnLeuGlyGlnIleValTyrValTyrAlaArgValSerTleAsnThrVal 200
DB 632 GGTCAATATTTCAGAAATTTGGGACGATGTTCTAGTGAGAGTTTCTGTGAACACAGCAAT 691
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgTyrHisGlyArgGlnTyr 220
DB 692 GTGACATCTGGGCTTCACATCTAGGAGTGCATCTACAGACAGCATGAGCGGCAAT 751
QY 221 IleProTlePheLysAspValTyrValTyrIleThrAspGlnIleProIlePheVal 240
DB 752 GTTCCATCTGCACAAAGTGTACGTACGTGGTAACAGATCAGATTCCTGTGTGTGTG 811

```

```

QY 241 ThrMetTyrGlyValLeuAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
DB 812 ACTATGTTCCAGAGAACCATGCAATTCATCCAGGAACCTTCCTCAAGATGTCGCC 871
QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
DB 872 ATTATGTTGATGCTCGATCATGCTTACGCCCTTCCTCAATATTCACATATAC 931
QY 281 TyrLysTTPAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
DB 932 TACAAGTGGAGCTTCGGGATAATCTAGGCTGTGTGTGTTCACCAATCATCTGTGAAT 991
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
DB 992 CACACGATGATGCTCATCGAAGCTTCAGCCTTAACCTCACTCTGNAAGCTCAGCACCA 1051
QY 321 GlyProCysProSerProThrProSerProSerSerThrSerProAspLeuSer 340
DB 1052 GNACTGTGTCCGCCGCCGCCACCCACCCAGACCT----- 1087
QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrLysSer 360
DB 1088 -----TCAAAACCCACCCCTCTTAGGACCTGCTGGTGACAAACCCC 1129
QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyPheArg 380
DB 1130 CTGCGCTGATGAGGATCTCTGATGAAGAACCTCCGATATACAGATATGCGCATTTCAA 1189
QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
DB 1190 GCCACATCAATTTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTC 1249
QY 401 ProLeuThrGlnGlnProAsnSerLeuMetAspPheIleValThrCysLysGly 420
DB 1250 CTGATCGGCTGCCATGGCTGAAAGCTCCCTATAGACTTGTGTGACTGCTGCCAAGG 1309
QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaAsn 440
DB 1310 AGCATTCGCCAGGAGGCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAAC 1369
QY 441 ArgValCysSerProValAlaValAspGluLeuGluCysLeuSerValArgAspAlaPhe 460
DB 1370 ACAATCTCGAGCCCTTGGATGTGATGATGATGTCTGCTGACTGTGAGACGACCTTC 1429
QY 461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
DB 1430 AATGCTGTGGAGCACTGTGTGATGCACTCCCTGGGAGATACAGACCTGCTGTGGT 1489
QY 481 ThrIleAlaLeuLeuThrTyrProGlyIleAspLeuGlySerProLeuArgThrValAsp 500
DB 1490 ACGAGACACCTGATTTCTCTGCTGACAGAGACCCAGCCCTGCTTTAGAGTGGCAAC 1549
QY 501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520
DB 1550 ACTGCCCTGATCTCGTGGCTTTGGCTTGGCCATTTGTGACTGTGTGTGCTGTGGT 1609
QY 521 TyrIleValHisIleValThrTyrLysProIleGlyAsnCysThrArgAsnValValIysGly 540
DB 1610 TACAAAAACACAGAAATACACCAATAGAAATAGTCTGGGAATGTGTGAGAAC 1669
QY 541 LysCylLeuSerValPheLeuSerHisIleLysAlaProPheSerArgGlyAspArgGlu 560
DB 1670 AAAGGCTGAGTGTCTTCTCAACGCTCAAAAGCGTGTTCTTCCTCCGSGGAACAGGAA 1729
QY 561 LysAspProLeuLeuGlnAspLys 568
DB 1730 AAGGTCGCTGCTACTCAAAACCAA 1753

```

RESULT 3
US-08-594-031-101
; Sequence 101, Application US/08594031
; Patent No. 5783182

GENERAL INFORMATION:
 APPLICANT: THOMPSON, Timothy C.
 TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 NUMBER OF SEQUENCES: 175
 CORRESPONDENCE ADDRESS:
 BAKER & BOTT, L.L.P.
 ONE WASHINGTON AVENUE, SUITE 1200
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20004-2400
 COMPUTER READABLE FORM:
 MODIFIED FROM: Diskette
 COMPUTER: IBM
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/594,031
 FILING DATE: 30-JAN-1996
 PRIORITY DATE: 30-JAN-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/006,838
 FILING DATE: 16-NOV-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Remenick, James
 REGISTRATION NUMBER: 36,902
 TELEPHONE: 202-639-7700
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-639-7700
 TELEFAX: 202-639-7890
 TELEX:
 INFORMATION FOR SEQ ID NO: 101:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2659
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 HYPOTHETICAL: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 US-08-594-031-101
 Alignment Scores:
 Seq. No.: 1,246-203
 Score: 2163.00
 Percent Similarity: 82.57%
 Best Local Similarity: 69.37%
 Query Match: 70.66%
 DB: 1

US-09-943-075a-2 (1-572) x US-08-594-031-101 (1-2659)
 QY 1 MetGluSerLeuGlyValLeuValPheLeuAlaGlyLeuProLeuIn 20
 DB 92 ATGGAGTCTCTACTACTTCTCGGGATTTCTGCTCGCTGGCTGCAAGATTGCCACTTGAT 151
 QY 21 AlaAlaLysArgAspValLeuGlyHisGluGlnTrpProAspHisMetArgLys 40
 DB 152 GCGCAACAGATTCATGATGCTCGGCATGAGGAGGCTCTCTCTATGAGGAG 211
 QY 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGlnTrpAspGluGlnLeuTrpPro 60
 DB 212 CACATCAATTAATAGGCTGCTCTCTGATGAAATAGCATGGAATGAAATCTTACCCA 271
 QY 61 ValTrpArgArgGlyGlyArgTrpLysAspSerTrpGluGlyValGlnAla 80
 DB 272 GTTGTGGAGGCGGAGAGCATAGGTGGAAATCTCTGGAGGAGGAGGCTGTGTGGAGG 331
 QY 81 AlaLeuTrpSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
 DB 332 GTCTGTGACGAGTACTACACAGCCCTGCTGGGCTGCAATATACATTTGCGGTGACCTG 391

QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTrpGluArgAsnCys 120
 DB 392 ATATCTCTCTAGATGCCAAAGAGAGATGCCAATGCCAATCTCTATGAGAGAACTGC 451
 QY 121 ArgSerAspLeuGluAlaSerAspProTrpValTrpAspGlnThrPheGlyAlaAsp 140
 DB 452 ACAAATGAGGCTGGTTTCTCTGCTGATCATATGTTTACAACTGGACAGCATGCTCAGAG 511
 QY 141 AspGluAspTrpLysAspAsnThrSerGlnGlyGlnIleLeuArgPheProAspGlyLys 160
 DB 512 GACAGTGGAGGGAAGAAATGCCAGCCGCAAGCCATCATACCTCTCTCCCTGATGGGAA 571
 QY 161 ProPheProArgProLysGlyValGlyLysTrpAspPheValTrpValPheHisThrLeu 180
 DB 572 CCTTTTCTCCACCAACCCGCGGAGGAAGATGGAATTTTCATCTAGCTCTTCCACCACTT 631
 QY 181 GlyGlnTrpPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
 DB 632 GTTCAGTATTTCCCAAAATGGACAGCTGTTTCAGTGAGAGTTCGTGGAACACAGCCAT 691
 QY 201 LeuThrValGlyProGlnIleValMetGluValIleValPheArgHisGlyArgAlaTrp 220
 DB 692 GTGACACTGTGGCTCAACTCATGAACTGACTGCTACAGAAGACATGGAGGGCATAT 751
 QY 221 IleProLysSerIleValLysAspValTrpValIleThrAspGlnIleProIlePheVal 240
 DB 752 GTTCCATGCGCAAGTGAAGATGTGTACGTGGTACAGATCAGATTTCTCTGTTGTG 811
 QY 241 ThrMetTrpGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
 DB 812 ACTATGTTCCAGAGAACTGGAATTCACGACGAACCTCTCTAAGATCTCCGC 871
 QY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTrpSerAlaIleSer 280
 DB 872 ATATGTTTGTGCTCTCATTCATGCTAGCTAGCCACTTCTCAATTATTTACCATTAAC 931
 QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300
 DB 932 TCAAGTGGAGCTTCGGGATTAATACGTGGCTTTTGTTCACCAATCATACTGTGAAT 951
 QY 301 HisThrTrpValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 DB 952 CACAGTATGCTCAATGAACTTCAGGCTTAACTCAGTGTGAAGCTCGACACCA 1051
 QY 321 GlyProCysProPheProThrProSerProSerSerSerThrSerProAlaSer 340
 DB 1052 GGACCTTTGTCGCCGCCGCCACCAACCCAGACT 387
 QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTrpLysSer 360
 DB 1088 -----TCAAAACCCACCTCTTTAGGCTCTGCTGGTGCACACCC 1129
 QY 361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTrpGlyTrpPheArg 380
 DB 1130 CTGAGCTCAGTAGGATTCCTCATGAAATGCCAATTAACAGATATGCCACTTCAA 1189
 QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
 DB 1190 GCGCACTCATCATGTAGAGGAGTCTAGAGTAACTATCATCTCAATCCAGACAGCTC 1249
 QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
 DB 1250 CTGATGCGCTGGCCATGGGCTGAAAGCTCCTTAATAGACTTTGCTGCTGACCTGCCAAGG 1309
 QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
 DB 1310 AGCATTCGACGAGGCTGTACCACTCATTTGACCCCACTTCGAGATCACCCAGAAC 1369
 QY 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460
 DB 1370 ACATCTGACGCCCTGTGGATGGATGGATGCTCTGCTGACTGTGAGAGCAACCTTC 1429

```
OY 461 AsnGlySerGlyThrTyrcysValasnPheThrLeuGlyAspAlaSerLeuAlaLeu 480
Db 1430 AYTGGGTCTGGGAGCTACTGTGTGACCTTCACCTCGGGGGATGACACAGCGCTGCTCTC 1489
OY 481 ThrSerAlaLeuSerLeuSerProGlyLysAspLeuGlySerProLeuAspThrValAsn 500
Db 1490 ACGAGCACCTGATTTCTGCTCTGACAGAGACCCAGCCTCGCCCTTAAAGGATGCAAC 1549
OY 501 GlyValLeuLeuSerIleGlyCysAlaLeuMetPheValThrMetValThrIleLeuLeu 520
Db 1590 ACTGCGCTTGATCTCCGTGCTGCTGGCCATATTGTCACATGATCTCCCTCTGGTG 1609
OY 521 TyrTyrlsValIleThrTyrlsPheTrileGlyAsnGlyThrArgAsnValValIleGly 540
Db 1610 TACAAAACACACAGCATACAAACCCAAATAGAAATATGCTCGGAATGTGGTCAGAGC 1669
OY 541 LysGlyLeuSerValPheLeuSerHisAlaIysAlaProPheSerArgGlyAspArgGlu 560
Db 1670 AAGAGCTTGATGCTTCTTCACGCTGCNAAGCCGTGCTTCCCGGGAACACAGAA 1729
OY 561 LysAspProLeuLeuIleAsnLys 568
Db 1730 AAGATCGCCTACTCAAAACCAA 1753

RESULT 4
US-08-594-031-113
US-08-594-031-113 Application US/08594031
Patent No. 5783182
GENERAL INFORMATION:
APPLICANT: THOMPSON, Timothy C.
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESS: BAKER & BOTT, L.L.P.
STREET: 1295 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2400
COMPUTER: IBM Compatible
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
PRIORITY DATE: 10-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remelick, James
STRATIFICATION NUMBER: 36,902
PREFERENCE TO FILE: 10-1446-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
SEQUENCE: 1056
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGIN: 1
ORIGINAL SOURCE:
US-08-594-031-113
Alignment Scores:
```

```
Pred. NO.: 6,27e-84 Length: 620
Score: 941.00 Matches: 183
Percent Similarity: 91.75% Conservative: 6
Best Local Similarity: 88.83% Mismatches: 15
Query Match: 30.74% Indels: 3
DB: 1 Gaps: 1
```

US-09-943-075A-2 (1-572) x US-08-594-031-113 (1-620)

```
OY 274 LeuAsnTySerAlaIleSerTyLysTrpAsnPheGlyAspAsnThrGlyLeuPheVal 293
Db 2 CTCACGACCTCGCCATTTCTCCACAGATGAACCTTGGGACACACACTCGCCCTGTTTGC 61
OY 294 SerAsnAsnHisThrLeuAsnHisThrTyValLeuAsnGlyThrPheAsnPheAsnLeu 313
Db 62 TCACACATCACTTGAATCACACTTATGTGCTCAATGGAACCTTCACACCTTAACCTC 121
OY 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSer-----ProSer 331
Db 122 ACCGTGCAACCTGACATGCGCGGCCACTGCCCTCCCTCGCTTCGACTTCGCCCTCCA 181
OY 332 SerSerThrProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351
Db 182 CTTCAACTCGCCCTCACCCTCGCCCTCACTCTGCCACATTTATCAACACCTAGCCCC 241
OY 352 SerLeuMetProThrGlyTyrlsSerMetGluLeuSerAspIleSerAsnGluAsnGly 371
Db 242 TCTTTAATGCTCTACTGTTACAAATCCATGAGAGCTGAGTGACATTTCCAAATGAATGC 301
OY 372 ArgIleAsnAsnGlyTyrlsGlyTyrlsPheAsnGlnThrIleThrIleValAspGlyIleLeuGlu 391
Db 302 CGAATAACAGATATGGCTACTTCAGAGCCACCATCACAAATTTAGAGGGATCTGGAA 361
OY 392 ValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnProAsnSerLeu 411
Db 362 GTCAGCATCTCGAGTAGAGTCTCCCTACGCCACACCGCAGCTCCGACTCCCTG 421
OY 412 MetAspPheIleValThrCysIysGlyIleAlaThrProThrGluAlaCysThrIleIleSer 431
Db 422 ATGACTTCTACTGTACCTGCNAAGGGGCCACCCCATGGAACCTGTAGATCATCTCC 481
OY 432 AspProThrCysGlnIleAlaAsnArgValCysSerProValAlaValAspGluLeu 451
Db 482 GACGCCACTGCCAGATCGCCAGACGGSTCTGCAGCCCTGTGGCTGTGGATGGCTG 541
OY 452 CysLeuLeuSerValArgArgAlaPheAsnGlySerGlyThrTyrcysValAsnPheThr 471
Db 542 TGC-G-CTGCTGTGAGAGAGCTTCATGTCGCTGGCAGCTACTGTGTGAATTTCACT 600
OY 472 LeuGlyAspAlaSer 477
Db 601 CTGGGAGATGATGCAAGC 618
```

RESULT 5

```
US-08-594-031-105/c
US-08-594-031-105/c Application US/08594031
Patent No. 5783182
GENERAL INFORMATION:
APPLICANT: THOMPSON, Timothy C.
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESS: BAKER & BOTT, L.L.P.
STREET: 1295 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2400
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
```

2 CURRENT APPLICATION DATA:
 3 APPLICATION NUMBER: US/08/594,031
 4 FILING DATE: 30-JAN-1996
 5 CLASSIFICATION: 35
 6 PATENT NO. 5763182
 7 PRIOR APPLICATION DATA:
 8 APPLICATION NUMBER: 60/006,838
 9 FILING DATE: 16-NOV-1995
 10 ATTORNEY/AGENT INFORMATION:
 11 NAME: Remenick, James
 12 REGISTRATION NUMBER: 36,902
 13 TELECOMMUNICATION INFORMATION:
 14 TELEPHONE: 202-639-7700
 15 TELEFAX: 202-639-7890
 16

2 INFORMATION FOR SEQ ID NO: 105:

3 SEQUENCE CHARACTERISTICS:
 4 LENGTH: 4,226-63
 5 TYPE: nucleic acid
 6 STRANDEDNESS: single
 7 TOPOLOGY: linear
 8 MOLECULE TYPE: cDNA
 9 HYPOTHETICAL: NO
 10 ANTI-SENSE: NO
 11 PRIMARY SOURCE: cDNA
 12 ORIGINAL SOURCE:
 13

2 US-08-594-031-105

Alignment Scores:

Pred. No.: 4,226-63 Length: 524
 Score: 729.00 Matches: 150
 Percent Similarity: 86.21% Mismatches: 18
 Query Match: 23.82% Indels: 2
 DB: 1 Gaps: 0

US-09-943-075A-2 (1-572) x US-08-594-031-105 (1-524)

Qy 389 TleGlValValMetIleGlnValAlaAspValProIleProThrLeuGlnProAsp 408
 Db 522 ATCTCGAAGTCAGCATGTCAGATAGCATGTCCCATGCCACACCGAGCCTGCC 463
 Qy 409 AsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCysThr 428
 Db 462 AATCCCTGATGAGCTTCATCTGACCTCGCAAGAGGGCCACCCCATGGAAGCCTGTAGC 403
 Qy 429 IleIleSerAspProThrCysGlnIleAlaIleAsnArgValCysSerProValAlaVal 448
 Db 402 ATCATCTCCGACCCACCTGCCATGCCATGCCAGACCGGGTCTGCAGCCCTGTGGCTGTG 343
 Qy 449 AspGluLeuCysLeuSerLeuValArgArgAlaPheAsnGlySerGlyThrTyrcysVal 468
 Db 342 GATGGGCTGTGGTG-CGTCTGTGAGAGAGCTTCATGGGTGGCCACTACTGTGG 284
 Qy 469 AsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIlePro 488
 Db 283 ANTTCACCTTGGGATGATCGAAGCCTGCCCTCCACACGACCCCTGCTCTATCCCT 224
 Qy 489 GlyLysAspLeuGlySerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCys 508
 Db 223 GGCACAGGCC-AGACTCCCTCGAGGCGATGATGTCTTATCTTCATCGGTCTC 155
 Qy 509 LeuAlaMetPheValThrMetValThrIleLeuLeuTyrlsLysHisLysThrTyrls 528
 Db 164 CTGGCTGTGCTGTGCACCATCTGTACCATCTCTGCTGACAAAACACACAGCGGTACAG 105
 Qy 529 ProIleGlyAsnCysThrArgAsnValValLysGlyLysGlyIleuSerValPheLeuSer 548
 Db 104 CCAATGGACATGCCCCAGGACACGATCAAGGCGAGGGCTGATGCTTCTCTCTCT 45
 Qy 549 HisAlaLysAlaProPheSerArgGlyAspArgGluLysAsp 562
 Db 44 CAGCGAAAGAGCCCGTCTCTCCGAGGACGACGAGGAGGAT 3

2 RESULT 6
 3 US-08-594-031-115
 4 Application US/08594031
 5 Filing Date: 30-JAN-1996
 6 Patent No. 5763182
 7 GENERAL INFORMATION:
 8 APPLICANT: THOMPSON, Timothy C.
 9 TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 10 NUMBER OF SEQUENCES: 175
 11 CORRESPONDENCE ADDRESS:
 12 JAMES REMENICK, JR.
 13 STEELES, 1299 Pennsylvania Avenue, N.W.
 14 CITY: Washington
 15 STATE: DC
 16 COUNTRY: USA
 17 ZIP: 20004-2400
 18 COMPUTER READABLE FORM:
 19 REGISTRATION NUMBER: 36,902
 20 TELECOMMUNICATION INFORMATION:
 21 TELEPHONE: 202-639-7700
 22 TELEFAX: 202-639-7890
 23

2 INFORMATION FOR SEQ ID NO: 115:
 3 SEQUENCE CHARACTERISTICS:
 4 LENGTH: 5,476-62
 5 TYPE: nucleic acid
 6 STRANDEDNESS: single
 7 TOPOLOGY: linear
 8 MOLECULE TYPE: cDNA
 9 HYPOTHETICAL: NO
 10 ANTI-SENSE: NO
 11 PRIMARY SOURCE: cDNA
 12 ORIGINAL SOURCE:
 13

2 US-08-594-031-115
 3 Application US/08594031
 4 Filing Date: 30-JAN-1996
 5 Patent No. 5763182
 6 GENERAL INFORMATION:
 7 APPLICANT: THOMPSON, Timothy C.
 8 TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 9 NUMBER OF SEQUENCES: 175
 10 CORRESPONDENCE ADDRESS:
 11 JAMES REMENICK, JR.
 12 STEELES, 1299 Pennsylvania Avenue, N.W.
 13 CITY: Washington
 14 STATE: DC
 15 COUNTRY: USA
 16 ZIP: 20004-2400
 17 COMPUTER READABLE FORM:
 18 REGISTRATION NUMBER: 36,902
 19 TELECOMMUNICATION INFORMATION:
 20 TELEPHONE: 202-639-7700
 21 TELEFAX: 202-639-7890
 22

2 INFORMATION FOR SEQ ID NO: 115:

3 SEQUENCE CHARACTERISTICS:
 4 LENGTH: 5,476-62
 5 TYPE: nucleic acid
 6 STRANDEDNESS: single
 7 TOPOLOGY: linear
 8 MOLECULE TYPE: cDNA
 9 HYPOTHETICAL: NO
 10 ANTI-SENSE: NO
 11 PRIMARY SOURCE: cDNA
 12 ORIGINAL SOURCE:
 13

2 US-08-594-031-115

Alignment Scores:

Pred. No.: 5,476-62 Length: 473
 Score: 717.00 Matches: 436
 Percent Similarity: 80.53% Mismatches: 63
 Query Match: 23.42% Indels: 2
 DB: 1 Gaps: 1

US-08-943-075A-2 (1-572) x US-08-594-031-115 (1-473)

Qy 274 LeuAsnThrSerAlaTleSerThrValProPheAsnThrGlyLeuPheVal 293
 Db 274 CTCACGACTGCGCATTTCTACAGTGGACTTTGGGGACACACCTGGCTGTGTG 61
 Qy 294 SerAsnAsnHisThrLeuAsnHisThrTyrlsValLeuAsnGlyThrPheAsnPheAsnLeu 313
 Db 62 TCCAAATACACACTTTGAATACACTTATGTCTCAATGGAACCTTCAACCTTAACTC 121
 Qy 314 ThrValGlnThrAlaValProGlyProCysProSerProThrProSer-----ProSer 331
 Db 122 ACCCTGCAATGCAAGTGGCCGGGACATGCCCTCCCTTCGCTTCGCTCGCCTCA 181
 Qy 332 SerSerThrSerProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351
 Db 332 SerSerThrSerProSerProAlaSerSerProSerProThrLeuSerThrProSerPro 351

```

Db 182 CTTTCACTCCGCCCTCACTCCGCCCTCGCCCTCGCCACATATTACACACCTAGCCGC 241
Qy 352 SerLeuMetProThrGlyTyrLysSerMetGluLeuSerAspIleSerAsnGluAsnCys 371
Db 242 TCTTTATTCCTTACTGGTTTAAATATCCATGGAGCTGAGTGACATTTCCATATGAACATCC 301
Qy 372 ArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValLeuAspGlyLysGlu 391
Db 302 CGAATAAACAGATATGGCTACTTCAGAGCCACCATCAGCATGTGTAGAGGGGATCTCGMA 361
Qy 392 ValAsnIleIleValAlaAspValProIleProThrLeuGlnProAsnSerLeu 411
Db 362 GTACATCATCATAGATACAGATATGTCATCGCCACCGCCAGCTGCCACATCCCTG 421
Qy 412 MetAspPheIleValThrCysTyrGlyAlaThrProThrGluAlaCysThr 428
Db 422 ATGGAGCTTCACTGTGACCTGCATGAAGGGGCCACCCCATGGAAGCCTGTACG 472

RESULTS
US-09-056-105-1
: Sequence 1, Application US/09056105
: Patent No. 6287569
: GENERAL INFORMATION:
: APPLICANT: KIPPS, THOMAS J.
: APPLICANT: KU, YUNOI
: TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
: PROCESSING
: FILE REFERENCE: 233/7 NUMBER: 105
: CURRENT APPLICATION NUMBER: US/09/056.105
: EARLIER FILING DATE: 1998-04-06
: EARLIER APPLICATION NUMBER: 60/043,467
: NUMBER OF SEQ ID NOS: 35
: NUMBER OF FASTSEQ For Windows Version 3.0
: SEQ ID NO: 1
: LENGTH: 2130
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-056-105-1

Alignment Scores:
Pred. No.: 4,55e-52 Length: 2130
Score: 627.50 Matches: 176
Percent Similarity: 41.85% Conservative: 96
Best Local Similarity: 27.08% Mismatches: 205
Query Match: 20.50% Indels: 173
Gaps: 19

US-09-943-075a-2 (1-572) x US-09-056-105-1 (1-2130)

Qy 7 ValLeuValPheLeuLeuAlaGlyLeuProLeuGluAlaAlaLys-----Arg 24
Db 46 CTCTCTCAATTTGGCTGTATAGTGTCTTGGCTGTGGGGGCTACAAAAGTACCACGA 105
Qy 25 PheArgValLeuGlyLysGluIleTyrProAspHisMetArgGluAsnGluLeu 44
Db 106 AACCAGAGCTGGCTTGGT-----GTCTCAAGGCCAATC 138

Qy 45 ArgGlyTrpSerSerAspGluAsnGluIleTrpAspGluLeuTyrProValTrpArgArg 64
Db 139 AGA-----ACCAAGCCTGGACAGCGCATGTATCCAGATGGACGA---180
Qy 65 GluGluGlyArgTrpLysAspSerTrpGluGlyValGlyValGluAlaLeuThrSer 84
Db 181 ---GAAGCCAGAGACTGACTGCTGGAGAGTGGTCAAGTGTCCCTCAAGGTCAAT 237
Qy 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
Db 238 GATGGGCTCACTGATGAGTAAGCAATGGCAATGGCTCTCTATGCTTCAAGCTTCCCTGGA 297
Qy 105 CysAsnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
Db 298 AGCCAAAGAGTATTGCCAGATGGCCAGTTATGTGGTGGTCAACATACCATCATCAATG 357

```

```

124 LeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAspAspGluAsp 143
358 AGCAGAGTGTGGGGAGGACAGCAGTGTATCCCGAGGAACT-----GACGATGCC--- 408
144 TrpGluAspAsnThrSerGlnGlyIleAsnIleLeuArgPheProAspGlyLysProPhePro 163
409 -----TGCNCTTCCCTGTGGTGTGAACCTTGCCCA 438
164 ArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisIleThrLeuGlyIleTyr 183
439 TCTGGCTCTTGGCTCTCAGAAGAAGCTTCTTTATGTCTGGGAAGACCTGGGGCCAATAC 498
184 PheGlnLysLeuGlyIleCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
499 TGGCAAGTCTTAGGGGGCCAGCTGTGGGCTGAGCATTTGGGACAGCGGCGCAATGCTG 558
204 GlyProIleValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222
559 GGCACACACACCATGSAAGTGACTGTCTACCATTCGCGGGGATCCGGAGCTATGTGCT 618
222 IleSerLysValLysAspValTyrValIleThrAspGluIleProIlePheValThrMet 242
619 CTTCGTCATTCAGCTCAGCTCTTCACCATTTACTACCATTTGCTTCTCCCTGAGCGTG 678
243 TyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
679 TCCAGTTCGGGGCTTGGATGGAGGACCAAGCAGCTCTCGAAGAACTAGCTCTGAC 738
263 PheAspValLeuIleHisAsnProSerHisPheLeuAsnTyrSerAlaIleSerThrLys 282
739 TTTGGCTCTCAGCTCCATGACCCAGCTGGCTATCTGTGGCTGAAGCTGACCTCTCTAC 798
283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisIleThrLeuAsnHisThr 302
799 TGGACCTTTGGAGCACTGTGTGAACCTGTATCTCTGGGCACTTGTGTGCTACTACT 858
303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnIleAlaValPro----- 320
859 TACCTGGAGCTGGCCAGCTACCTGCCAGGTGGTCTGCAGCTGCCATCTCTCTAC 918
321 -----GlyProCysPro----- 324
919 TCTCTGGCTCTCCAGTTCCAGGCACACAGATGGCAGCAGCACTGCAGAGGCC 978
324 ----- 324
979 CTTAACACACAGCTGGCCAGTCCCTACTACAGAACTGTGGGTGACTACACTGGTGCAG 1038
325 SerProThrProSerProSerSerSerThr----- 334
1039 GCGCAACTGCAGAGCCCTCTGGAACCACTATCTGTGCGAGTGGCAACCACTCAAGTCATA 1098
335 SerProSerProAlaSerSerProSerProThrThrLeuSerThrProSerProSerLeuMet 354
1099 AGCAGCTCACTGTGCACTGCCAGTCCAGACACAGAGACACAGGTATGACACTGAGAGGTG 1158
355 Pro-----ThrGly 357
1159 CCATCTTTCAGAGGTCTAGGTTACCACACTGGCAGAGATGTCAACTCCAGAGGCTACAGT 1218
358 TyrLysSerMetGluLeuSerAspIle----- 366
1219 ATGACACCTTGCANAGTATCATTTGTGGTGTCTTCTGGAACCACTGCACAGGTACCA 1278
366 ----- 366
1279 ACTAGAGTGGGTGGAGACACAGCTAGAGAGCTACTATCTCTGAGCCTGAAGGTCCA 1338
367 -----SerAsnGlu----- 369
1339 GATGCCAGCTCAATCATGTCTTACGAGAGATTATACAGGTTCCTGCGGGCCCTCTGCTGGAT 1398

```

:	TELEX:	421792	
:	INFORMATION FOR SEQ ID NO:	26:	
:	SEQUENCE CHARACTERISTICS:		
:	LENGTH:	2172	
:	TYPE:	nucleotide	
:	STRANDEDNESS:	Double	
:	TOPOLOGY:	Unknown	
:	MOLECULE TYPE:	cDNA	
:	US-08-417-174-26		
Alignment Scores:			
Pred. No.:	5,27e-52	Length:	2172
Score:	627.00	Matches:	174
Percent Similarity:	41.03%	Conservative:	96
Best Local Similarity:	26.44%	Mismatches:	206
Query Match:	20.48%	Indels:	182
DB:	2	Gaps:	19
US-09-943-075A-2 (1-572) x US-08-417-174-26 (1-2172)			
Qy	17	LeuPProteinGluAlaIalysargPheArgAspVal.....	-----28
Db	12	TTACCAATCGGCAGCCGGAGAACAACATGCATGTGTCTAAAAGATSCCTTTTCAT	71
Qy	29LeuGIHylGluInTyPro.....	---Asp 36
Db	72	TTTGCGTGTGATAGCTGTTTGTGCTGTGGGGGTACAAAAAGTACCACGAACCCAGAC	131
Qy	37	HslMetArgGluAsnAsnInLeuArgGlyTrpSerAspGluAsnGluIrrAspGlu	56
Db	132	TGCGCTGGTGTCTCAGGCCACTCAGA.....	---ACCAAGCTCTGGACMGK 176
Qy	57	GluLeuYrProValIrrPAArgGlyGluGVargITryAspSerITrgGluGlyGly	76
Db	177	CACCTGTATCCACGAGTGGACA.....GAACCCGACAGACTTGACTGCTGGAGAGTGCT	230

[illegible]

```

Db 732 TTCTCGAGAAATCAGCCTCTGACCTTTGGCCCTCAGCTGCATGACCCGACCTGGCTATCTG 791
Qy 275 AsnTyrSerAlaIleSerTyrTyrTrpAsnPheGlyAspAsnThrGlyLeuPheValSer 294
Db 792 CGTACAGCTGACCTCTCTCACTGACCTGGGACTTTGGAGACAGTACTGAACCCCTGACTCT 851
Qy 295 AsnAsnHisThrLeuAsnHisThrTyrValValLeuAsnGlyThrPheAsnLeuThr 314
Db 852 CGGGCACTTGTGGTGGCTACTTACTTACTTGGAGCTGGCCAGCTCACTGCCAGTGTC 911
Qy 315 ValGlnThrAlaValPro-----GlyProCysPro----- 324
Db 912 CTCAGAGCTGCATCTCTCTCACTCTGCTGGCTCTCCCGCACTGTCAGGACACAGAT 971
Qy 324 ----- 324
Db 972 GGGCACAGGCCAAGCTGACAGAGCCCTTAACACACAGCTGGCCAGTGCCTACTACAGAA 1031
Qy 325 ----- 334
Db 1032 GTTGTGGTACTACACTGCTGACGGGGCCCACTGCAGAGCGCTCTGGAACCACTCTGTG 1091
Qy 335 ----- 335
Db 1092 CAGGTGCCAACCACTGAAGTCAATGAGACACTGACACTGTGGAGTGCCTCACTGAGAGC 1151
Qy 347 SerThrProSerProSerLeuMetPro----- 355
Db 1152 ACAGGTATGAGACCTGAGAAGTGCCAGTTTCAGAGTCAATGGTACCACACTGGCAGAG 1211
Qy 356 ----- 366
Db 1212 ATGTCACTCCAGAGCTACAGATGTGACACTGCAGAGGATCATTTGTGGTCTTCT 1271
Qy 366 ----- 366
Db 1272 GGAAACACCTGCACAGGTAAACAACCTACAGAGTGGGTGGAGACACACAGCTAGAGCTA 1331
Qy 367 ----- 369
Db 1332 CTTACTCTGAGCTGACAGTCCAGAGTCCAGCTCACTCACTACCTACGGAAGATTACA 1391
Qy 369 ----- 369
Db 1392 GGTTCCTCTGGCCCTCTGCTGATGGTACACACCACTTAAGCTGTGGTAGAGCAAGTC 1451
Qy 370 -----AsnCysArgGlyIleAsnArgTyrGlyPheArgAlaThrIleThrIleValAsp 387
Db 1452 CCCTGGATGTTGTCTGTATGATGATATGCTTCTTTCCTGCACCTCGCATTTGTCAG 1511
Qy 388 GlyIleLeuIleValAsnIleIleGlnValAlaAspValProIleProThrLeuIlePro 407
Db 1512 GTATTGAAAGTCCGACAGTCTCTCCAGGCT-----GTGCGCTGGGTGAGGG 1559
Qy 408 AspAsnSerLeuMetAspPheThrIleValThrCysGlyAlaThrProThrGlyIleCys 427
Db 1560 GAT-----GCATTGGTGACTGTGCTCTGCAAGCGGGCTGCCCAAGCAAGCTGC 1613
Qy 428 ThrIleIleSerAspProThrCysGlnIleAlaIleAsnArgValCysSerProValAla 1673
Db 1614 ATGGAGATCTCATCGCCAGGTGCCAGGCTCCCGCCAGCGCTGTGCCAGCTGTGCTA 1733
Qy 448 ValAspGluLeuSerLeuSerValIleArgAlaPhe---AsnGlySerGlyThrTyr 466
Db 1674 CCGACGCCAGCTGCCAGCTGGTGTCTACACAGATACCTGAAGGGTGGCTCGGGGACATAC 1733
Qy 467 CysValAsnPheThrIleGlyAspAspAlaSerIleAlaLeuThrSerAlaLeuIleSer 486
Db 1734 TGCTCAATCTGCTCTGCTGATACCAACAGCTGGCAGTGGTGCACCCAGCTTATC 1793
Qy 487 IleProGlyIleAspLeuGly---SerProLeuArgThrValAsnGlyIleValLeuIleSer 505

```

```

Db 1794 ATCCGCTGCTCAAGAAGCAGGCTTGGCAGGTTCCGCTGATCGTGGGCATCTTG----- 1847
Qy 506 ILeGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrGlyLys 523
Db 1848 -----CTGGTGTGTAAGCTGTGCTCTGCTCTGATATATAGGCGC 1892
RESUL 9
US-08-231-565A-26
Sequence 26, Application US/08231565A
Patent No. 5874560
GENERAL INFORMATION:
APPLICANT: KANAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A. MELANOMA ANTIGENS AND
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 145 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
OPERATING SYSTEM: IBM PC COMPATIBLE
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,565A
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
AUTHOR: KANAKAMI, YUTAKA; ROSENBERG,
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH 2172
TYPE: nucleotide
STRANDINESS: Double
ORIENTATION: 5' to 3'
MOLECULE TYPE: CDNA
US-08-231-565A-26
Alignment Scores:
Pred. No.: 5,27e-52 Length: 2172
Score: 627.00 Matches: 964
Score: 627.00 Conserved: 964
Best Local Similarity: 26.44% Mismatches: 206
Query Match: 20.48% Indels: 182
Gaps: 19
US-09-943-075A-2 (1-572) x US-08-231-565A-26 (1-2172)
Qy 17 LeuProIleGlnAlaAlaIleValArgPheArgAspVal----- 28
Db 12 TTACCAATCGCAGCGGGAAGACACATGATCTGTGCTAAAAGATGCTTCTTCAT 71
Qy 29 -----LeuGlyHisGlnIlePro----- 36
Db 72 TTGGCTGTGATAGTGTCTTTCCTGCTGGGGCTACAAAGGTACCCAGAACCGAGC 131
Qy 37 HisValArgGluAsnAsnGlnLeuArgGlyTyrPheSerAspGluAsnGlyIleTyrPAspGlu 56
Db 132 TGGCTTGGTGTCTCAAGCACTCAGA-----ACCAAGCGCTGGAACAG 176
Qy 57 GlnLeuTyrProValTrpArgArgGlyGlyLeuIleArgTrpLysAspSerTrpLuglyGly 76

```


; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/231,565
 ; FILING DATE: 22-APR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CAROL M. GRUPPI
 ; REGISTRATION NUMBER: 37,341
 ; TELEPHONE: (212) 751-6840
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 751-6849
 ; TELEFAX: (212) 751-6849
 ; TELETYPE: 421792
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; TYPE: nucleotide
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Unknown
 ; MOLECULE TYPE: cDNA
 ; US-09-007-961-26

Alignment Scores:
 Pred. No.: 5,27e-52 Length: 2172
 Score: 627.00 Matches: 174
 Percent Similarity: 41.03% Conserved: 96
 Best Local Similarity: 26.44% Mismatches: 206
 Query Match: 20.48% Indels: 182
 Gaps: 19

US-09-943-075A-2 (1-573) x US-09-007-961-26 (1-2172)

QY 17 LeuProLeuGlnAlaAlaIysArgPheArgAspVal-----28
 DB 12 TTNCAATCGCGACGGGAGAACACATGGATCTGGTGTAAAGATGCCCTTCTCAT 71
 QY 29 -----LeuGlyVHISLugInPro-----Asp 36
 DB 72 TTGGCTGCTAGTGTCTTGGCTGTGGGGCTACAAAGTACCAGAACCCAGGAC 131
 QY 37 HSMetArgGluAAsnArgGlnLeuArgGlyTrpSerSerAspGluAAsnGluTrpAspGlu 56
 DB 132 TGCTTGGTGTCTCAAGCGACACTCAGA-----ACCAGAGCTGGACAGG 176
 QY 57 GlnLeuTrpProValTrpArgGlyGluGlyArgTrpLysAspSerTrpLugLugLys 76
 DB 177 CAGCTGTATCCAGCTGTGGAC-----CAGCGCCAGAGAGTTGACTGCTGGAGAGTGT 230
 QY 77 ArgValGlnAlaAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96
 DB 231 CANGTGTCCCTCAAGTATGATATGATGGCTTACACTGTGTGGTACAAAGCTCTTTC 290
 QY 97 ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnLysValTrp 116
 DB 291 TCTATTGCTGTGACTCTCCCTGGAGCCAAAGAGTATTGCCAGATGGCGAGCTATCTGG 350
 QY 117 GluArgan---CysArgSerAspGluLeuAlaSerAspProTyrValTrpAsnTrp 135
 DB 351 GTCACAACTTACCATCATCATGGGACGAGCTGGGGAGACGACGCGATATCCCGAC 410
 QY 136 ThrThrValAlaAspAspGluAspTrpGluAspAsnThrSerGlnGlnHisLeuArg 155
 DB 411 GAACT-----GACGATGCC-----TGATC 431
 QY 156 PheProAspGlyLysProPheProArgProHisGlyArgLysTrpAsnPheValTrp 175
 DB 432 TTCCCTGATGCTGGACCTTGCCCATCTGTGGCTCTTGGTCTCAGAGAGAGCTTTGTAT 491
 QY 176 ValPheHisThrLeuGlyInTrpPheGlnLysLeuGlyGlnCysAsrAlaArgValSer 195
 DB 492 GTCGTGGAGAGCTGGGCGCAATGCTGCGCATCTGGCATTTCTAGGGGGCGGCTGCTGGCTGAG 251
 QY 196 IleAsnThrValAlaAsnThrValGlyProGlnValMetCysValIleValPheArg 447
 DB 552 ATTTGGGACGAGGCGCATGCTGGGCGCACACACACACATGGAGTGTGCTTACCATCGC 611

QY 216 HisGly---ArgAlaTrpIleProIleSerLysValLysAspValTrpValIleThrAsp 234
 DB 612 CGGGAGTCCGGAGCTATGTGCTCTTGCATCTCCAGCTCAGCTTCCACCATTAAGTGAC 671
 QY 235 GlnIleProIlePheValThrMetTrpGlnLysAsnAspArgAsnSerSerAspLysThr 254
 DB 672 CAGTGGCTTTCTCCCTGAGCTGAGCTGTCCGATTTGGGGGCTTGGATGGAGGACACACAC 731
 QY 255 PheLeuArgAspLeuProIlePheAspValIleuIleHisAspProSerHisPheLeu 774
 DB 732 TTCCTGAGAAATCAGGCTCTGACCTTTGGCCTCCAGCTCCATGACCCAGTGGCTATCTTG 791
 QY 275 AsnTrpSerIleIleSerTrpLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSer 294
 DB 792 SCN3AAGCTGACCTCTCTACCTGGGACTTTGGAGACAGTAGTGGAAACCTGATCTCT 851
 QY 295 AsnAsnHisThrLeuAsnHisIleThrValLeuAsnGlyThrPheAsnPheAsnLeuThr 314
 DB 852 CGGCGACCTGTGGTCTACCTACTTACCTTGGAGCTTGCCCGACCTGCGCCAGGTGGTC 911
 QY 315 ValGlnThrAlaValPro-----GlyProCysPro-----324
 DB 912 CTGAGGCTGCCATCTCTCTCACCTCTGTGGCTCTCCCTCCAGTTCAGGCGACACAGAT 971
 QY 324 -----324
 DB 972 GGGCAGGCGCACTCGAGAGGCCCTTACACACAGCTGGCCAGTGGCTACTACAGAA 1031
 QY 325 -----SerProThrProSerProSerSerSerThr-----334
 DB 1032 GTTCTGGTACTACACTGGTGGAGGCCCACTGCAGAGGCTCTGGAAACACCATCTGTG 1091
 QY 335 -----SerProSerProAlaSerSerProSerProThrLeu 346
 DB 1092 CAGGTGGCAACCACTCGAGTCAATACCACTGCACTGTGCGAGTGGCACTGCGAGAGC 1151
 QY 347 SerIleProSerProSerLeuMetPro-----355
 DB 1152 ACAGGTATGACACCTGAGAGGTGGCGAGTTTCAGAGGTGATGGGTACCACTGGCAGAG 1211
 QY 356 -----ThrGlyTrpLysSerMetGluLeuSerAspIle-----366
 DB 1212 ATGTCAACTCCAGAGCTACAGGTATGACACATGCGAGAGTCAATGATGTGGTCTTCT 1271
 QY 366 -----366
 DB 1272 GGAAACAGCTGGCAAGGTATACACTACAGGTGGTGGAGACACAGCTAGAGAGTA 1331
 QY 367 -----SerAsnGlu-----369
 DB 1332 CCTATGCTGAGCCTGAGAGTCCAGATGCCAGTCAATCATGTCTACGGAAGATTACA 1391
 QY 369 -----369
 DB 1392 GTTGTGCTGGGCCCTGCTGGATGCTTACGCGCACCTTAAGGCTGGTGAAGACAGTTC 1451
 QY 370 -----AsnCysArgLysAsnArgTrpGlyTrpPheArgAlaThrIleThrIleValAsp 387
 DB 1452 CCCTCGGATGTGTGTGTATCATGATGGTCTTTCTCCGTCACCTGCACTGGATGTCAC 1511
 QY 388 GlyIleLeuGluValAsnIleGlnValAlaAspValProIleProThrLeuGlnPro 407
 DB 1512 GTTATGAAGTGGCGAGTCTCTCGAGCT-----GTGCGTGGGTGGGG 1559
 QY 408 AspAsnLeuMetMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCys 427
 DB 1560 GAT-----GCATTTGAGCTGACTGTGCTGCGAAGGGGCGTGGCCGAAGAGCTGCG 1613
 QY 428 ThrIleIleSerAspProThrCysGlnIleAlaGlnAsnArgValCysSerProValAla 447
 DB 1614 ATGGAGATCTCATCGCAGGCTGCCAGGCCCTCTGCCAGCGGCTGTGCCAGCTGTGCTA 1673

Oy 448 ValaspGluLeuLysLeuSerValArgAlaPhe---AsnGlySerGlyThr 466
 Db 1674 CCAGCCAGCCAGCTGCCAGCTGGTCTTCGACCCAGATACGAAGAGGTGGCTGGGAGCATAC 1733
 Oy 467 CysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuLeuSer 486
 Db 1734 TGCTCATGTGCTCTGGCTGTATACCAACAGCCTGGCAGTGGTGCAGCCAGCTTATC 1793
 Oy 487 IleProGlyLysAspLeuGly---SerProLeuArgThrValAsnGlyValLeuLeuSer 505
 Db 1794 ATGGCTGGTCAAGAGAGGAGCTGGGAGAGGTTCGGCTGATCTGGGAGCATCTTG--- 1847
 Oy 506 IleGlyLysLeuAlaMetPheValThrMetValThrIleLeuLeuThrValLys 523
 Db 1848 ---CTGCTGTGATGGCTGGTGGCTGGCTGGCATCTCTGATATATAGGCG 1892
 RESULT 11
 US-09-267-439-26
 : Sequence 26 Application US/09267439
 : No. 629
 : GENERAL INFORMATION:
 : APPLICANT: KAKAMAKI, YUTAKA; ROSENBERG,
 : APPLICANT: STEVEN A.
 : TITLE OF INVENTION: MELANOMA ANTIGENS AND
 : TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
 : TITLE OF INVENTION: METHODS
 : NUMBER OF PAGES: 26
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 : STREET: 345 PARK AVENUE
 : CITY: NEW YORK
 : STATE: NEW YORK
 : COUNTRY: USA
 : ZIP: 10155
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: FLOPPY DISK
 : COMPUTER: IBM PC COMPATIBLE
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: ASCII DATA:
 : CHARACTER SET: ASCII
 : APPLICATION NUMBER: US/09/267,439
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/417,174
 : FILING DATE: 05-APR-1995
 : APPLICATION NUMBER: US/08/231,565
 : FILING DATE: 05-APR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: CAROL M. GRUPPI
 : REGISTRATION NUMBER: 37,341
 : REFERENCE/DOCKET NUMBER: 2026-4124US1
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (212) 759-4800
 : TELEFAX: (212) 751-6849
 : TELETYPE: 431792
 : INFORMATION FOR SEQ ID NO: 26:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2172
 : TYPE: nucleotide
 : STRANDEDNESS: Double
 : TOPOLOGY: Unknown
 : MOLECULE TYPE: cDNA
 US-09-267-439-26

Alignment Scores:
 Pred. No.: 5,27e-52 Length: 2172
 Score: 627.00 Matches: 174
 Percent Similarity: 41.03% Conservative: 96
 Best Local Similarity: 26.48% Identical: 146
 Query Match: 20.48% Gaps: 19

US-09-943-075a-2 (1-572) x US-09-267-439-26 (1-2172)

Oy 17 LeuProLeuGlnAlaLeuAlaLysArgPheArgAspVal- 28
 Db 12 TTACCATATGCGACCGGAGAACACAACTAGATCTGGTGTGCTGTAAAGATGCTTCTTCATP 71
 Oy 29 ---LeuGlyIleGluGlyPro- 36
 Db 72 TTGGCTGTGATAGTCTCTTGGCTGGTGGGGGCTACAAAGATACCCAGAACACGAGC 131
 Oy 37 HlMetArgLysAsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGlu 56
 Db 132 TGGCTGTGGTCTTCAGGCACTCAGA- 176
 Oy 57 GlnLeuTrpProValTrpArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGly 76
 Db 177 CAGCTGTATCATCAGAGTGACA- 230
 Oy 77 ArgValGlnAlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPhe 96
 Db 231 CAGTGTCCCTCAGGTCACTATGATATGAGGCTACACTGATGGTGCATATGCCCTCTCT 290
 Oy 97 ValValAsnLeuValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTrp 116
 Db 291 TCTATTCCTGCTCAACTCCCTCGAAGCCAAAGGTATTGCCAGATGGGAGGTATTCTGG 350
 Oy 117 GlnArgAsn---CysArgSerAspLeuGluLeuAlaSerAspProTrpValTrpAsnTrp 135
 Db 351 CTCACAAATACCATCATCATCAATGGAGCGAGGTGGGGAGGACGACGAGCTGTATCCCGAG 410
 Oy 136 ThrThrGlyAlaAspAspGluAspTrpGluAspAsnThrSerGlnGlyGlnIleLeuArg 155
 Db 411 GAAACT---GAGCATGCC- 431
 Oy 156 PheProAspGlyLysTrpPheProArgProHlSerGlyArgLysLysTrpAsnPheValTrp 175
 Db 432 TTCCCTGATGTGGAGCTTGCCTATCTGGCTCTTGGCTCTCAGAGAGAGAGCTTGTATTAT 491
 Oy 176 ValPheHlStrLeuGlyIleTrpPheGlnLysLeuGlyIleCysSerAlaArgValSer 195
 Db 492 GTCTGGAGAGACTGGGCGCATACCTACCTGGCATTTCTAGGGGGCCGAGTGTGGGTGAGG 551
 Oy 196 IleAsnThrValAsnLeuThrValGlyProGlnValMetGluValIleValPheArgArg 215
 Db 552 ATTTGGAGAGCGAGGCGCAATGCTGGGCACACACACCATCGAAGTGTCTCATCATCGC 611
 Oy 216 HlGly---ArgAlaTrpIleProIleSerTrpValLysAspValTrpValIleThrAsp 234
 Db 612 CGSGATCCCGAGCTATGTGCTCTTGGCTATTCAGCTCAGGCTTCACCATTAATGAC 671
 Oy 235 GlnIleProIlePheValThrMetTrpGlnLysAsnAspArgAsnSerSerAspGluThr 254
 Db 672 CAGTGTCTTTCTCCGAGGCTGTCCAGTTGGGGGCGCTTGGATGAGGGAGAGGACAGCAG 731
 Oy 255 PheLeuArgAspLeuProIlePhePheAspValLeuIleHlAspProSerHlSerPheLeu 274
 Db 732 TTCTTGAAATCAGCTCTGACCTTGGCTCCCTCCAGCTCCATCATGACCCCATGCTGTCTG 791
 Oy 275 AsnTrpSerIleSerTrpLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSer 294
 Db 792 CTTGAGAGTCACTCTCTACCTGAGCTTTGGAGACGAGTGTGAGACCCCTGATCTCT 851
 Oy 295 AsnAsnHlStrThrLeuAsnHlStrThrValValLeuAsnGlyThrPheAsnPheAsnLeuThr 914
 Db 852 CGGCGCACTTCTGTGTCACTACTTACTCTGAGGCTGGCCGCTGACCTGCCAGGTGCTC 1031
 Oy 315 ValGlnThrAlaValPro---GlyProCysPro- 324
 Db 912 CTCGAGGCTGCCATTCCTCTACCTCTCTCTGGCTCTCTCCCTGGTTCAGGACGACACAGAT 971
 Oy 324 --- 324
 Db 972 GGGCAGAGGCCAACTGCAGAGGCCCTTAACACACAGCTGGCCAAAGTGCTACTACAGAA 1031

```

OY 325 -----SerProthrProSerSerSerThr----- 334
Db 1032 GTTGTGGTACTACACTGCTGAGCGCCGACAGAGCCCTCTGAGACACTGTGTG 1091
OY 335 -----SerProthrProSerSerSerProthrLeu 346
Db 1092 CAGGTGCCACCACTGAAGTCATPACACTGCACTCTGACAGTCCACCTGACAGC 1151
OY 347 SerThrProSerProSerLeuMetPro----- 355
Db 1152 ACAGATATGACCACTGAGAAGTGCCAGCTTTCAGAGGTCATGCTGGTACCACTGGCAGAG 1211
OY 356 -----ThrGlyTyrLysSerMetGluLeuSerAspIle----- 366
Db 1212 ATGTCMACTCCAGAGGCTACAGGTATGACACTGCCAGAGTATCAATTGTGTGCTTCT 1271
OY 366 -----SerArgLys----- 366
Db 1272 GMAACCACTGCTACAGGTATCAACTAGAGTGGTGGAGACCACTAGAGACTA 1331
OY 367 -----SerArgLys----- 369
Db 1332 CCTATCGCTGAGCTGGAAGTCCAGATGCCAGCTCAATCATGCTACGAAAGATTACA 1391
OY 369 -----SerArgLys----- 369
Db 1392 GGTTCCTGGCCCTCTGATGCTGACAGCACCCTTAAGCGGTGAGAGCAAGTC 1451
OY 370 -----AsnCysArgIleAsnArgTyrGlyTyrPheArgAlaThrIleThrIleValAsp 307
Db 1452 CCCTCGGATTTGCTCTGTATGATATGTTCTTTTCCGTCACCCCTGGAGATGTCCAG 1511
OY 388 GlyLeuGluValAsnIleIleGlnValAlaAspValProIleProThrLeuGlnPro 407
Db 1512 GGTATGMAAGTCCAGATCTCTGAGCT-----GTGCGCTCGGTGAGGG 1559
OY 408 AspAsnSerLeuMetAspPheIleValThrCysLysGlyAlaThrProThrGluAlaCys 427
Db 1560 GAT-----GCAITTAGCTGACTGTGCTGCGCAAGCGCGCTGCCAAGGAAGCTGCG 1613
OY 428 ThrIleIleSerAspProThrCysGlnIleAlaGlnAsnAspValIcySerProValAla 447
Db 1614 ATGGAGATCATGCGCAGGCTGCCAGCCCTGCCCGCGCTGTGCGAGCTGTGCTA 1673
OY 448 ValAspGluLeuCysLeuLeuSerValArgAlaPhe---AsnGlySerGlyThrTyr 466
Db 1674 CCAACCCAGCTGCCAGCTGTGCTGTGCGCCACACTACTGAGGCTGCTGGGCAATAC 1733
OY 467 CysValAspPheThrLeuGlyAspAspAlaSerLeuAlaLeuThrSerAlaLeuIleSer 486
Db 1734 TGCCTCAATGTGCTCTGGCTGATACCAACAGCTGTCACAGTGGTCAGCCAGCAGCTATTC 1793
OY 487 IleProGlyLysAspLeuGly---SerProIleuArgThrValAsnGlyValLeuIleSer 505
Db 1794 ATGCTGTGTCAGAAACAGCGCTGTGGCAGGTTCGCTGTGCTGTGGCCTCTTG 1847
OY 506 IleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeuTyrLysLys 523
Db 1848 -----CTGGGTGTGATGGCTGTGGTCTGCTGCTATATATAGGGC 1892

```

```

RESULT 12
US-09-594-031-92
: Sequence 92, Application US/08594031
: Patent No. 5783182
: GENERAL INFORMATION:
: APPLICANT: THOMPSON, Timothy C.
: TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
: OF DNA FROM TISSUE SAMPLES
: CORRESPONDENCE ADDRESS: 175
: ADDRESSEE: BAKER & BOTTS, L.L.P.
: STREET: 1299 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: DC

```

```

: COUNTRY: USA
: ZIP: 20004-2400
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: CURRENT APPLICATION DATA:
: CURRENT APPLICATION DATA: Version 1.5
: APPLICATION NUMBER: US/08/594,031
: FILING DATE: 30-JAN-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 90/006,838
: FILING DATE: 01-01-1993
: ATTORNEY AGENT INFORMATION:
: NAME: Remenick, James
: REGISTRATION NUMBER: 36,902
: REFERENCE/DOCKET NUMBER: 0A146-0110
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-639-7700
: TELEFAX: 202-639-7890
: INFORMATION FOR SEQ ID NO: 92:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 335 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: COORDINATE SYSTEM: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
: US-08-594-031-92
: Alignment Scores:
: Pred. No.: 3,648-43
: Score: 524.00
: Percent Similarity: 88.70%
: Best Local Similarity: 86.09%
: Query Match: 17.12%
: Models: 4
: Gaps: 1
: US-09-943-075a-2 (1-572) x US-08-594-031-92 (1-335)
OY 225 LysValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMetTyrGln 244
Db 1 AAGTGAAGAAGTGTATGTGTATACAGTACAGATCCCTGTATCTGTGACATGTCGCG 60
OY 245 LysAsnAspArgAsnIleThrPheLeuArgAspLeuProIlePhePheAsp 264
Db 61 AAGATGACAGGAGACTGTGTGATGAGATCTCTGCAGACACCTCCCATCTCTTCGAT 120
OY 265 ValLeuIleIleAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrIleTyrAsn 284
Db 121 GTCTCATCATGATCCAGCCACTTCTCTCCACGACTCTGCCATCTCTACACGTGCAAC 180
OY 285 PheClyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyrVal 304
Db 181 TTTCGGGACACACATGCGCTTTGTCTCCCAACATCACACTTTCATCATCATATGTG 320
OY 305 LeuAspGlyThrPheAsnPheAsnLeuThrValGlyThrAlaValProGlyProCysPro 344
Db 241 CTCATGGAACCTTCAACCTTAACCTCAACGCTCAACGCTCTGCCATCTCTACACGTGCAAC 300
OY 325 SerProThrProSerProSerSerThrSerProSerProAla 339
Db 301 CCCCCCTTGGCT-----TCGACTCCGCTCCACCTTCG 333

```

```

RESULT 13
US-08-594-031-94
: Sequence 94, Application US/08594031
: Patent No. 5783182

```

GENERAL INFORMATION:

APPLICANT: THOMPSON, Timothy C.
 TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 NUMBER OF SEQUENCES: 175
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BAKER & BOTTS, L.L.P.
 STREET: 1259 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: DC

COUNTRY: USA
 ZIP: 20004-2400
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/594,031
 FILING DATE: 30-JAN-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/006,838
 FILING DATE: 16-NOV-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Remedick, James
 REGISTRATION NUMBER: 36,902
 REFERENCE/DOCKET NUMBER: 0A146-0110
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-639-7700
 TELEFAX: 202-639-7890

INFORMATION FOR SEQ ID NO: 94:

SEQUENCE CHARACTERISTICS:
 LENGTH: 335
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 FRAGMENT TYPE: NO
 ORIGINAL SOURCE:
 US-08-594-031-94

Alignment Scores:

Pred. No.: 335
 Scores: 524.00
 Percent Similarity: 86.09%
 Best Local Similarity: 86.09%
 Query Match: 17.12%
 DB: 1

US-09-943-075a-2 (1-572) x US-08-594-031-94 (1-335)

Qy 225 LysValLysAspValThrValIleThrAspGlnIlePheValThrMetTyrGln 244
 Db 1 AAGTGGAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTGTCACCATGTGCCAG 60
 Qy 245 LysAsnAspAgaAsnSerSerAspGluThrPheLeuArgAspLeuPhePheAsp 264
 Db 61 AAGATGACAGAGAACTGTCTGATGAGATCTTCCTCAGAGACCTCCCATGCTTCGAT 120
 Qy 265 ValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrIleTyrAsn 284
 Db 121 GTCTTCATTCATGATCCAGCCACTCTCCTCAAGCACTCTGCCATTTCTTCAAGTGGAC 180
 Qy 285 PheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyrVal 304
 Db 181 TTGGGGACACACTGGCTGTGTGTCTCCACCACTCACTTGTATGCACACTTATGAG 240
 Qy 305 LeuAsnGlyThrPheAsnLeuThrValGlnThrAlaValProGlyProCysPro 324
 Db 241 CTCATATGGAACCTTCAACCTTAACCTTCAACCTGCAAGTGCAGTGCCTGGGCGCATGCCCT 300

Qy 325 SerProThrProSerProSerSerThrSerProSerProAla 339
 Db 301 CCCCTTGGCT-----TCGACTCGGCTCCCATCTTCC 333

RESULT 14

US-08-594-031-96
 : Sequence 96, Application US/08594031
 : Patent No. 5783182
 : GENERAL INFORMATION:
 : APPLICANT: THOMPSON, Timothy C.
 : TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 : NUMBER OF SEQUENCES: 175
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BAKER & BOTTS, L.L.P.
 : STREET: 1259 Pennsylvania Avenue, N.W.
 : CITY: Washington
 : STATE: DC

COUNTRY: USA
 ZIP: 20004-2400
 COMPUTER: IBM Compatible
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/594,031
 FILING DATE: 30-JAN-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/006,838
 FILING DATE: 16-NOV-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Remedick, James
 REGISTRATION NUMBER: 36,902
 REFERENCE/DOCKET NUMBER: 0A146-0110
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-639-7700
 TELEFAX: 202-639-7890

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:
 LENGTH: 335 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 US-08-594-031-96

Alignment Scores:
 Pred. No.: 335
 Scores: 524.00
 Percent Similarity: 88.70%
 Best Local Similarity: 86.09%
 Query Match: 17.12%
 DB: 1

US-09-943-075a-2 (1-572) x US-08-594-031-96 (1-335)

Qy 225 LysValLysAspValThrValIleThrAspGlnIlePheValThrMetTyrGln 244
 Db 1 AAGTGGAAGATGTGTATGTGATAACAGATCAGATCCCTGTATTGTCACCATGTGCCAG 60
 Qy 245 LysAsnAspAgaAsnSerSerAspGluThrPheLeuArgAspLeuPhePheAsp 264
 Db 61 AAGATGACAGAGAACTGTCTGATGAGATCTTCCTCAGAGACCTCCCATGCTTCGAT 120
 Qy 265 ValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrIleTyrAsn 284
 Db 121 GTCTTCATTCATGATCCAGCCACTCTCCTCAAGCACTCTGCCATTTCTTCAAGTGGAC 180

Db 121 GTCTCATTCATGATCCAGCCACTTCCTCAACGACTCGCCATTTCCTACAGTGGAC 180
 Oy 285 PheGlyAspAnthrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyVal 304
 Db 181 TTTGGGAGCAACACTGGCTGTTTGTCCTCCAAATCACTTTGAATCACACTTATGTG 240
 Oy 305 LeuAsnGlyThrPheAspPheAspLeuThrValGlyCysThrAlaValProGlyProCysPro 324
 Db 241 CTCATGGAACCTTCAACCTTAACCTCACTCGCTGCAACTGCAGTGGCCGGCCATGGCCCT 300
 Oy 325 SerProThrProSerProSerSerThrSerProSerProAla 339
 Db 301 CCCCTTCGCCCT-----TCGACTCCGCTCCACCTTCG 333
 RESULT 15
 : US-08-594-031-123
 : Sequence 123, Application US/08594031
 : Patent No. 5783182
 : GENERAL INFORMATION:
 : INVENTOR: ROBERT J. HARRISON, Timothy C.
 : TITLE: ANTIMETASTATIC METHOD FOR IDENTIFYING METASTATIC SEQUENCES
 : NUMBER OF SEQUENCES: 178
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BAKER & BOTTS, L.L.P.
 : STREET: 1299 pennsylvania Avenue, N.W.
 : CITY: Washington
 : STATE: DC
 : COUNTRY: USA
 : ZIP: 20004-2400
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq Version 1.5
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/594,031
 : FILING DATE: 30-JAN-1996
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/006,838
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Remenick, James
 : REGISTRATION NUMBER: 36,902
 : REFERENCE/DOCKET NUMBER: 0A146-0110
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-639-7700
 : TELEFAX: 202-639-7850
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 123:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 335 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : HYPOTHEetical: NO
 : ANTI-SENSE: NO
 : FRAGMENT TYPE:
 : ORIGINAL SOURCE:
 : US-08-594-031-123

Alignment Scores:
 Pred. No.: 3-64e-43 Length: 335
 Score: 524.00 Matches: 99
 Percent Similarity: 88.70%
 Best Local Similarity: 86.09%
 Query Match: 17.12%
 Indels: 4
 Gaps: 1
 DB:

US-09-943-075A-2 (1-572) x US-08-594-031-123 (1-335)

Oy 225 LysVallyAspValTyValIleThrAspGlnIleProIlePheValThrMetTyGln 244

Db 1 AAGGTGAAAGATGTTATGTGTATGATACAGTCAGATCCCTGTTCTGTGACCATGTCCTCG 60
 Oy 245 LysAsnAspAsgAsnSerSerAspGluThrPheLeuAlaGspLeuProIlePhePheAsp 264
 Db 61 AAGAAATGACAGGAACCTTGCTGTGATGAGATCTCTCTCAGAGACCTCCCATCTGCTTCGAT 120
 Oy 265 ValLeuIleHisAspProSerHisPheLeuAsnTySerAlaIleSerTyTyIleTyPheAsn 284
 Db 121 GTCTCATTCATGATCCAGCCACTTCCTCAACGACTTCGCCATTTCCTACAGTGGAC 180
 Oy 285 PheGlyAspAnthrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThrTyVal 304
 Db 181 TTTGGGAGCAACACTGGCCCTTTTGCTCCAAATCACTTTGAATCACACTTATGTG 240
 Oy 305 LeuAsnGlyThrPheAsnPheAsnLeuThrValGlyCysThrAlaValProGlyProCysPro 324
 Db 241 CTCATGGAACCTTCAACCTTAACCTCACTCGCTGCAACTGCAGTGGCCGGCCATGGCCCT 300
 Oy 325 SerProThrProSerProSerSerThrSerProSerProAla 339
 Db 301 CCCCTTCGCCCT-----TCGACTCCGCTCCACCTTCG 333

Search completed: April 9, 2003, 13:08:30
 Job time : 84 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 9, 2003, 07:59:06 : Search time 151 seconds
(without alignments)
3322.781 Million cell updates/sec

Title: US-09-943-075a-2

Perfect score: 2061

Sequence: 1 MESLQGLVFLAALGLPLQ.....PPSGDKREKLQKHPWL 572

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Delop 6.0, Delext 7.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -cpu=vlh

-db=/cgn2_1/US07POOL/US0943075/runat_07042003_155632_18004/app_query.fasta.1.711

-db=PublishedApplications_NA -qfmt=fastap -suffix=rnpb -minmatch=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BL0SUM62

-TRANS=human40.cdl -LIST=45 -DOCNALIGN=200 -THR_SCORE=pcr -THR_MAX=100

-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pl0 -NORM=ext -HEAPSIZ=500 -MINLEN=0

-NCHLEN=2000000000 -USER=US0943075 -CCN=0000000000 -37 -runat_07042003_155632_18004

-NCHLEN=2000000000 -USER=US0943075 -CCN=0000000000 -37 -runat_07042003_155632_18004

-DEV_TIMOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=0.5 -XGAPEXT=0.5 -FCGAPOP=6

-FCGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_NA:

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3061	100.0	2320	10	US-09-943-075a-1
2	2700	88.2	2213	10	US-09-943-075a-1
3	2669	88.2	2213	10	US-09-943-075a-1
4	2163	70.7	2069	9	US-09-545-978B-82

Sequence 41, Appl	9	US-10-227-884-41	2683	70.6	2162	5
Sequence 41, Appl	6	US-10-230-163-41	2162	70.6	2162	6
Sequence 41, Appl	7	US-10-218-631-41	2162	70.6	2162	7
Sequence 41, Appl	8	US-10-230-338-41	2162	70.6	2162	8
Sequence 41, Appl	9	US-10-230-338-41	2162	70.6	2162	9
Sequence 41, Appl	10	US-09-943-075a-1	2162	70.6	2162	10
Sequence 41, Appl	11	US-09-943-075a-1	2162	70.6	2162	11
Sequence 1, Appl1	9	US-10-039-272-1	2661	69.8	2136	9
Sequence 1, Appl1	12	US-09-812-238B-1	627.5	20.5	627.5	12
Sequence 1, Appl1	13	US-09-862-260A-1	627.5	20.5	627.5	13
Sequence 1, Appl1	14	US-10-047-539-3	622	20.3	622	14
Sequence 1, Appl1	15	US-10-047-539-3	620	20.3	620	15
Sequence 1, Appl1	16	US-09-960-352-7359	453	16.1	453	16
Sequence 1, Appl1	17	US-09-960-352-7359	453	16.1	453	17
Sequence 1, Appl1	18	US-09-910-941-210	445	14.5	445	18
Sequence 1, Appl1	19	US-09-864-761-1520	405.5	13.2	405.5	19
Sequence 1, Appl1	20	US-09-917-800A-139	299	9.8	299	20
Sequence 1, Appl1	21	US-09-864-761-8527	295.5	9.7	295.5	21
Sequence 1, Appl1	22	US-09-864-761-8527	295.5	9.7	295.5	22
Sequence 1, Appl1	23	US-09-864-761-8527	295.5	9.7	295.5	23
Sequence 1, Appl1	24	US-09-864-761-8527	295.5	9.7	295.5	24
Sequence 1, Appl1	25	US-09-925-301-7061	287	9.6	287	25
Sequence 1, Appl1	26	US-09-960-352-5490	257	8.4	257	26
Sequence 1, Appl1	27	US-09-962-598-240	165	5.4	165	27
Sequence 1, Appl1	28	US-09-989-293A-240	165	5.4	165	28
Sequence 1, Appl1	29	US-09-989-735-240	166	5.4	166	29
Sequence 1, Appl1	30	US-09-989-735-240	166	5.4	166	30
Sequence 1, Appl1	31	US-09-989-735-240	166	5.4	166	31
Sequence 1, Appl1	32	US-09-989-735-240	166	5.4	166	32
Sequence 1, Appl1	33	US-09-989-735-240	166	5.4	166	33
Sequence 1, Appl1	34	US-09-989-735-240	166	5.4	166	34
Sequence 1, Appl1	35	US-09-989-735-240	166	5.4	166	35
Sequence 1, Appl1	36	US-09-989-735-240	166	5.4	166	36
Sequence 1, Appl1	37	US-10-174-590-179	166	5.4	166	37
Sequence 1, Appl1	38	US-10-174-590-179	166	5.4	166	38
Sequence 1, Appl1	39	US-10-174-590-179	166	5.4	166	39
Sequence 1, Appl1	40	US-10-174-590-179	166	5.4	166	40
Sequence 1, Appl1	41	US-10-174-590-179	166	5.4	166	41
Sequence 1, Appl1	42	US-10-174-590-179	166	5.4	166	42
Sequence 1, Appl1	43	US-10-174-590-179	166	5.4	166	43
Sequence 1, Appl1	44	US-10-174-590-179	166	5.4	166	44
Sequence 1, Appl1	45	US-10-174-590-179	166	5.4	166	45

ALIGNMENTS

```
RESULT 1
US-09-943-075a-1
; Application US/09943075A
; Reference: US/09943075A
; GENERAL INFORMATION
; APPLICANT: Popoff, Steven F.
; APPLICANT: Safado, Faye F.
; APPLICANT: Owen, Thomas A.
; APPLICANT: Snock, Steven L.
; TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,
; TITLE OF INVENTION: Compositions and Methods of Stimulating Bone Differentiation
; REFERENCE: US/09943075A
; CURRENT APPLICATION NUMBER: US/09/943,075A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,006
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 2320
; TYPE: DNA
; ORGANISM: Rat osteoactivin
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(1833)
US-09-943-075a-1
```

Alignment Scores:

```

Pred. No.: 7,636-315 Length: 2320
Score: 4061.00
Matches: 572
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10
Gaps: 0

US-09-943-075a-2 (1-572) x US-09-943-075a-1 (1-2320)

QY 1 MetGluSerLeuCyGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20
DB 115 ATGGAAGTCTCTGCGGGGCTCTGGTATTTCTGCTGCTGCTGCGAGACTGCGGCTCG 174
QY 21 AlaAlaLysArgPheArgValLeuGlyHisGlnGlnTrpProAspHisMeLeArgGlu 40
DB 175 GGGGCAACACGGGTCTGGTGTGCTGGCGCATGAGCNGATCCGGATACATGAGGGAG 234
QY 41 AsnGlnLeuLeuGlyGlyTrpSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
DB 235 AACACCAATTAGTGGTGGTCTTCAGATGAAATGATGGATGGATGAACAGCTGTATCCA 254
QY 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpCylGlyArgValGlnAla 80
DB 295 GTGTGGAGAGGGAGAGGGAGAGATGAAGAGACTCTCGGGAGAGAGGGCGGTGTGAGCCA 354
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnLleThrPheValValAsnLeu 100
DB 355 GCCCTAACCAAGTATTCACCGGCTTGGTGGGTTCCTCAATATCACTTCTGTAGTGAACCTG 414
QY 101 ValPheProArgCysGlnLysAlaAsnGlyAsnLleValTrpGluArgAsnCys 120
DB 415 GTGTTCCTCCATGCTGAGAGAGAGATGCCACGCAATATCTGCTATGAGAGCAATGCT 474
QY 121 ArgSerAspLeuGluLeuAlaSerAspProTrpValTrpAsnTrpThrGlyValAsp 140
DB 475 AGAAGTCTATTGGAGCTGGCTTCTGACCGCTATGTCTTACAACTGGACACAGGGGAGAC 534
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyHisLeuArgPheProAspGlyLys 160
DB 535 GATGGAGCTGTGGAGACACACACACACACACACACACACACACACACACACACACAC 594
QY 161 ProPheProArgProHisGlyArgLysArgTrpAsnPheValTrpValPheHisThrLeu 180
DB 595 CCCTCTCCCTCCGCCCCACGGACGAGAAATGGAACTTCTGCTAGCTCTTCCACACACTT 654
QY 181 GlyGlnTrpPheGlnLysLeuGlyGlnCysSerAlaLysValSerLleAsnThrValAsn 200
DB 655 GTTCAGTATTTTCAAAAGCTGGGTGATGCTTCAGCAGAGTTCCTATATTAACAGTCAAC 714
QY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTrp 220
DB 715 TTGACAGTTGGCCCTCAGGTCTAGGAAGTATGATGCTTTTGGAGACAGCGCGGGCATAC 774
QY 221 IleProGlnSerTrpValLysAspValTrpValLleThrAspGlnLleProLlePheVal 240
DB 775 ATTCCATCTCCAAAGCTGAGTATGATGATACAGATCAGATCTCTATATCTGTTG 834
QY 241 ThrMetLysGlnLysAsnArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
DB 835 ACATGTACAGAGAAATGACGCAACTGCTGTCTGTAAGCACTCTCTGAGACACTCTCCC 894
QY 261 IlePhePheAspValLeuLleHisAspProSerHisPheAsnThrSerLleHisLeu 280
DB 895 ATTTCTCTGATCTCTCTCATTCAGATCCCATGCTATTTCTCTCACTACTCTGCTCATTC 954
QY 281 TyrLysTrpAsnPheCylAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300
DB 955 TCAAGTGGAACTTTGGGGACACACTGGGCTGTGTTGCTCCCAATCACTCACTTTGAAT 1014
QY 301 HisThrTrpValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
DB 1015 CACACGTATGTGTCAATTGAACCTTCAACTTTAACTTCACTGCGAACTGCACTGCGG 1074

```

```

QY 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340
DB 1075 GACCATGCCCCCTTACCCACACACTGCGCTTCTTCTGACCTTCTTCTGCGCTGCTTCT 1134
QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTrpLysSer 360
DB 1135 TCGCCTTCCACCACATATATCAACACACTAGTCCCTCTTAAATGCTTACTGCGTCAAAATCC 1194
QY 361 MetGluLeuSerAspLleSerAsnGlnAsnCysArgLleAsnArgTrpGlyTrpPheArg 380
DB 1195 ATGAGCTGTAGTGTATCTCCATGAACACTGCGGAATAAACATATGTTACTTCTCAGA 1254
QY 381 ArgThrLleThrLleValAlaCysGlyLleValLeuGluValAlaAspVal 1300
DB 1255 GCCACCATCAACAATGTAGATGAATCTCTAGAACTCAACATCATCCAGTAGCAGCATGTC 1314
QY 401 ProLleProThrLeuGlnProAspAsnSerLeuMetAspPheLleValThrCysLysGly 420
DB 1315 CCATCCCAACACATCCAGCTGTACACTCTGATGAGACTTCTATGTGTACTGTGCAAGGG 1374
QY 421 AlaThrProThrGluAlaCysThrLleLleSerAspProThrCysGlnLleAlaGlnAsn 440
DB 1375 GCCACTTCCACCGAGAGCTGTAGCATCATCTCTGACCCACCTGCCAGATGCCCAAGAAC 1434
QY 441 ArgValCysSerProGluValAlaAspLleCysCysLeuLeuSerValArgArgAlaPhe 460
DB 1435 AGSGTCTCCAGCCGGTGGCTGTGCTGATGAGTGTGCTCTCTGCTGAGGAGGCGCTTC 1494
QY 461 AsnGlySerGlyThrTrpCysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu 480
DB 1495 AATGGTCTCCGCGACGTACTGTGTGAATTTACTCTGGGAGAGCATGCCAAGCTGCGCTC 1554
QY 481 ThrSerAlaLeuLleSerLleAspGlyLysAspLeuGlySerProLeuArgThrValAsn 500
DB 1555 ACACGCGCCCTGATCTCTATCTCTGSCAAAGACCTAGCTCCCTCTCTGAGACAGTGAAT 1614
QY 501 GlyValLeuLleSerLleGlyCysLeuAlaMetPheValThrMetValThrLleLeuLeu 520
DB 1615 GGTGTCTCTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1674
QY 521 TyrLysLysHisLysThrTrpLysProLleGlyAsnCysThrArgAsnValValLysGly 540
DB 1675 TACAAAAACACAGACGTACAAAGCCAAATAGGAACTGCACCAAGCAACCTGGTCAAGGCC 1734
QY 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgLysAspArgGlu 560
DB 1735 AAGGCGCTGTAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1794
QY 561 LysAspProLeuLeuGluAspLysProTrpMetLeu 572
DB 1795 AAGGATCTACTGCTCCAGCAAGCAAGCATGGATGCTC 1830

RESULT 2
US-09-943-075a-27 Application US/09823038A
US-09-943-075a-27 Application US/09823038A
; Patent No. US20020058335A1
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevlin
; APPLICANT: Omrust, Rene
; APPLICANT: Mulder, Gerd
; APPLICANT: Mulder, Gerd
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; FILE REFERENCE: 11000.1037C3
; CURRENT APPLICATION NUMBER: US/09/823-038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOURCE: GenBank
; SEQ ID NO 27
; LENGTH: 2213
; TYPE: DNA

```


ORGANISM: Mouse
US-09-823-038A-27

Alignment Scores: 1,51e-276 Length: 2213
Sred. No.: 27,400
Parent Similarity: 92,29%
Best Local Similarity: 88,79%
Query Match: 88,21%
DB: 11
Gaps: 1

US-09-943-075A-2 (1-572) x US-09-823-038A-27 (1-2213)

Qy 1 MetGluSerLeuCyGluValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
Db 23 ATGGAAGATCTCTCGGGGCTCTGGGATTTCTGCTGGCTCGAGGACTGCTTCCAG 82
Qy 21 AlaAlaValArgPheArgValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
Db 83 CTCTGACAGGAGATTCGTGATGTCTGGCCATGGAACAGTAATCCCATCATCATGAGAG 142
Qy 41 AsnAsnGlnLeuATGCTGCTPSeSerAspGluAsnGlnTyrPaspGluGlnLeuTyrPro 60
Db 143 CACACCAATATACCTGGCTGGTCTTCGGATGAAATGAATGGATGAACACTGTATATCCA 202
Qy 61 ValTTPArgGArgGlyGluGlyArgTTPylsAspSerTPpGluGlyArgValGlnAla 80
Db 203 GTGTGAGAGGGAG 262
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnLleThrPheValValAsnLeu 100
Db 263 GTCTCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 332
Qy 101 ValPheProArgCysGlnGlyGluGluAlaSerAlaAsnGlyAsnLleValTyrGluArgAsnCys 120
Db 323 GTTGTCCCGACAGATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 382
Qy 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTTPThrThrGlyAlaAsp 140
Db 383 AGGATGATTTGGGACTGACCTCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 442
Qy 141 AspGluAspTPpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 443 GATGTGACTGGGAGATGGCAGCAGCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAG 502
Qy 161 ProPheProArgProHisGlyArgLysLysTPpAsnPheValTyrValPheHisThrLeu 180
Db 503 CCTCTCTCTGCGCCCGCTGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 562
Qy 181 GlyGlnTyrPheGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 200
Db 563 GGGCAGTATTTCCAAAACCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 622
Qy 201 LeuThrValGlyProGlnValMetGluValLleValPheArgArgHisGlyArgAlaTyr 220
Db 623 TTGACATCTGGCCCTCATGGCTATGGAGAGAGACTGTCTTTCGAGATCATGGCGGGCAT 682
Qy 221 IleProHisSerLysValLysAspValTyrValLleThrAspGlnLleProIlePheVal 240
Db 683 ATTCCCATCTCGAAGGTCAAAGATGTATGTCTGATACAGATCATGATCCCTGTATTCG 742
Qy 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
Db 743 ACCGTGTCCCGAGAGAGATGACAGGACTGTGTGTGATGATGATCTCTCTCGAGACCTCC 802
Qy 261 IlePhePheAspValLeuLleHisAspProSerHisPheLeuAsnTyrSerAlaLleSer 280
Db 803 ATGGTCTTGGATGTCTCATTCATGATTCACGAGGACCTTCTCCACAGGACTCTGCCATTTC 862
Qy 281 TyrLysTPpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnLleThrLeuAsn 300
Db 863 TACAGTGGAGACTTGGGAGAACACTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 922

Qy 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Db 923 CACACTTATGTGCTCAATGGAACCTTCAACCTTAACCTCAACCTGCAACCTGCACTGCC 982
Qy 321 GlyProCysProSerProThrProSerProSerProSerProSerProSerProSerPro 338
Db 983 GGGCATATGCTTCCCTCTGGCTCTGCACTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1042
Qy 339 AlaSerProSerProThrLeuSerThrProSerProSerProSerProSerProSerPro 358
Db 1043 CGGCGCTTCCCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1102
Qy 359 LysSerMetGluGluSerAspLleSerAsnGluLeuAsnGlyArgLysArgTyrGlyTyr 378
Db 1103 AAATCCATGGAGCTGAGTGTGATCTTCAATGAAACTCCGGAATTAACAGATATGGCTAC 1162
Qy 379 PheArgAlaThrLleThrLleValAspLysLleLeuGluValAsnLleLleGlnValAla 398
Db 1163 TTTAGAGGCGACCATCAATATGTAGAGGGATCTTGGAACTGAGCATCATGCAATGCA 1222
Qy 399 AspValProLleProThrLeuGlnLleProAspSerSerLeuMetAspPheLleValThrCys 418
Db 1223 GATTTCCCATGCTCCACACGCGCTGCCAATCTCCCTGATGGACTTCACTGTGACCTGC 1282
Qy 419 LysGlyAlaThrProThrGluAlaCysThrLleLleSerAspProThrCysGlnLleAla 438
Db 1283 AAAGGGCGACCCCATGAGAGGCTGTACATCATCTCCGACCCACCTCCCATGCTGCC 1342
Qy 439 GlnAspArgValCysSerProValAlaValAspGluLeuCysLeuSerValArgArg 458
Db 1343 CAGAACCGGGCTGTGCAACGCTGTGGCTGTGGATGGCTGTGGCTGTGGCTGTGGAGAA 1402
Qy 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478
Db 1403 GCTTCATATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1462
Qy 479 AlaLeuThrSerAlaLeuLleSerLleProGlyLysAspLeuGlySerProLeuLeuArg 498
Db 1463 GCGCTCAGCAGACCTGTATCTATCTCTGCAAGACGACGACTCCCTCTGAGAGCA 1522
Qy 499 ValAsnGlyValLeuLleSerLleGlyCysLeuAlaMetPheValThrMetValThrIle 518
Db 1523 GTGATGGTGCTCTGATCTCCATTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1582
Qy 519 LeuLeuTyrLysLysHisLysTyrTyrProLleGlyAsnAsnCysThrArgAsnValVal 538
Db 1583 TTGCTGTGCAAAAACAGAGCGCTACAGCCCAATAGAGAACTGCCCGACAGACGGCT 1642
Qy 539 LysGlyLysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgLysAsp 558
Db 1643 AAGGGCAGAGGCTGTGTCTCTTCCAGCCAGCGAGAGGCGGCTCTCTCCGAGGAGAC 1702
Qy 559 ArgGlyLysAspProLeuGlnAsnLysAspLysPro 569
Db 1703 CAGGAGAGAGATCATCTGCTCCAGGAGAGCA 1735

RESULT 3
US-09-943-075A-7
Sequence 7: Application US/09943075A
Patent No. US20020151486A1
GENERAL INFORMATION:
APPLICANT: Safado, Feyer F.
APPLICANT: Popoff, Steven N.
APPLICANT: Owen, Thomas A.
INVENTOR: Owen, Thomas A.
TITLE OF INVENTION: Osteostatin Protein and Nucleic Acids Encoding the Same.
FILE OF INVENTION: Compositions and Methods of Stimulating Bone Differentiation
FILE REFERENCE: 71369,262
CURRENT APPLICATION NUMBER: US/09/943,075A
PRIOR APPLICATION NUMBER: 2001-08-30
PRIOR FILING DATE: 2000-08-30
NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 725

; LENGTH: 1725

; ORGANISM: Mouse

US-09-943-075A-7

Alignment Scores:

Pred. No.: 3,42e-276

Score: 2695.00

Identical Similarity: 42

Best Local Similarity: 86.73%

Query Match: 88.04%

DB: 1

Gaps: 10

Length: 1725

Matches: 507

Mismatch: 168

Indels: 2

Gaps: 1

US-09-943-075A-2 (1-572) x US-09-943-075A-7 (1-1725)

QY 0Y 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20
 Db 1 ATGGAAAGTCTCTCGGGGTCCTGGGATTTCTGCTGCTGGCTGAGACAGCTCTCCAG 60
 QY 21 AlaAlaLysArgPheAetGaspValLeuGlyHisGlnGlnTyrProAspHisMetArgGlu 40
 Db 61 GCTGCCAAGCATTCGTGATGTCCTGGGCCATGAGACATATCCGATCATCATCAGAGAG 120
 QY 41 AenAsnGlnLeuArgCylProSerSerAspGluAsnGlnLupAspGluGlnLeuTyrPro 60
 Db 121 CACACCAATTACGTGGTCTTCGGATGAATGATGGATGGACACCTCTATCCA 180
 QY 61 ValTrpArgGlyGlyGlyArgTyrPlyAspSerTrpGluGlyValArgValGlnAla 80
 Db 181 GTGTGGAGGGGAGGACGAGAGGTGGAGGACTCTCTGGGAGGAGGCGCTGTGAGCCA 240
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnLleThrPheValValAsnLeu 300
 Db 241 GTCTCTGACCACTGACTCACCAGCTCTGTGGGTTCACATATACACTTTTGTGGTGAACCTG 300
 QY 101 ValPheProArgCysGlnTyrGluAspAlaAsnGlyAsnLleValTyrGluArgGlnCys 120
 Db 301 GTGTTCCTCCAGTCCGAGAGAGATGCTATGGCAATTCGTCTATGAGAGACATCC 360
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140
 Db 361 AGGAATGATTGGGACTGACATCTGACCTGCTACACTGACCTGACGGGCGAGAT 420
 QY 141 AspGluAsnTrpGluAspAsnTrpSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
 Db 421 GATGGTACTGGGAAGATGGCACCCGAGGACGAGCATCTCAGTCTCCGACAGGAGG 480
 QY 161 ProPheProArgProHisGlyArgTyrMetGluValPheValPheHisThrLeu 180
 Db 481 CCCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 540
 QY 181 GlyGlnTyrPheGlnTyrLeuGlyGlnCysSerAlaHisValSerLleAsnThrValAsn 200
 Db 541 GGCCAGTATTTCCAAAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
 QY 201 LeuThrValGlyProGlnValMetGluValLleValPheArgArgHisGlyArgAlaTyr 220
 Db 601 TTGACAGCTGGCCCTCAGGCTCAGGAAGTACTGCTCTTCGAAGATAGCGCGGCGCATC 660
 QY 221 IleProLleSerLysValLysAspValTyrValLleThrAspGlnLleProLlePheVal 240
 Db 661 ATTCCTCATCGAGGAGAGATGCTGTTGATACAGATCAGATCCCTGATGATCGTG 720
 QY 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
 Db 721 ACCATGTCCGAGAGATGACAGAACTGCTGTGATGAGATCTCTCTCAGAGACCTCCCC 780
 QY 261 IlePhePheAspValLeuLleHisAspProSerHisPheLeuAsnTyrSerLleIleSer 280
 Db 781 ATCTGCTTCGATGCTCTCATTCATGATCCAGCACTCTCTCAACGACTCTGCCATTCCTC 840

QY 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
 Db 841 TACAGTGGAACTTTGGGACACACATGGCTGTGTGCTCCACACACACACTTTGAT 900
 QY 301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 Db 901 CACACTTATGCTCAATGGAACCTTCAACCTTAACCTCAGCTGCAAACTGGACGTGCC 960
 QY 321 GlyProCysProSerProThrProSer-----ProSerSerSerThrSerProSerPro 338
 Db 961 GGCCATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1020
 QY 339 AlaSerSerProSerProThrThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 358
 Db 1021 TGGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1080
 QY 359 LysSerMetGluLeuSerAspLleSerAsnGluAsnCysArgLleAsnArgTyrGlyTyr 378
 Db 1081 AATTCATGGAGCTGGAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 379 PheArgAlaThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 398
 Db 1141 TTGAGAGCACCATTACAAATTTAGAGAGGATCTCGAAGTCAGCATCATCATCATCATCAT 1200
 QY 399 AspValProLleProThrProThrProAsnSerLeuMetAspPheLleValPheCys 418
 Db 1201 GATGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1260
 QY 419 LysGlyAlaThrProThrGluAlaCysThrLleLleSerAspProThrCysGlnLleAla 438
 Db 1261 AAGGGGGCCACCCCATGGAAGCTGTGACATCATCTCCGACCCACCTGCCAGATGCC 1320
 QY 439 GluAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgPro 458
 Db 1321 CAGAACGGCTCGAGCCCTGTGGCTGTGGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 QY 459 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 478
 Db 1381 CCCTTCATGGTGTGGCCCTACTGTGATGATTCATCTGGAGATGATGCCAGCTG 1440
 QY 479 AlaLeuThrSerAlaLeuLleSerLleProGlyLysAspLeuGlySerProLeuArgThr 498
 Db 1441 CCCTCCACCAAGCACCCTGATCTCTATCCCTGCCAAAGACCCAGACTCCCTCTGAGACA 1500
 QY 499 ValAsnGlyValLeuLleSerLleGlyCysLeuAlaMetPheValThrMetValThrIle 518
 Db 1501 GTGATGT 1560
 QY 519 LeuLeuTyrLysHisLysHisLysHisLysHisLysHisLysHisLysHisLysHisLys 538
 Db 1561 TTGCTGTACAAAAACCAAGGCTGACAGCCCAATAGGAACCTGCCAGGAACAGCGTC 1620
 QY 539 LysGlyLysGlyLeuSerValPheLeuSerHisLalLysAlaProPheSerArgGlyAsp 558
 Db 1621 AAGGGCGAGGGGCTGAGGTGCTCTGATGATCCACCCCAAGCCGCTTCTCCGAGGAGAC 1680
 QY 559 ArgGluLysAspProLleSerPro-----ProSerSerSerThrSerProSerPro 569
 Db 1681 CAGGAGAGAGGATCCATTTGCTCAGAGACAGCA 1713

RESULT 4

US-09-525-978b-82

; Application No. US2003004972A1

; Publication No. US2003004972A1

; GENERAL INFORMATION:

; APPLICANT: Murray, Richard

; APPLICANT: Caras, Ingrid W.

; APPLICANT: Hevesi, Peter

; APPLICANT: Willson, Keith

; TITLE OF INVENTION: METHODS OF DIAGNOSING MACROPHAGE DEVELOPMENT

; TITLE OF INVENTION: RELATED DISORDERS, COMPOSITIONS, AND METHODS OF

; TITLE OF INVENTION: SCREENING FOR MACROPHAGE DEVELOPMENT MODULATORS

; FILE REFERENCE: A-67413-1/DJB/JJD

1 CURRENT APPLICATION NUMBER: US/09/525, 978B
2 CURRENT FILING DATE: 2000-03-15
3 PRIOR APPLICATION NUMBER: US9N 60/124, 530
4 PRIOR FILING DATE: 1999-03-15
5 NUMBER OF SEQ ID NOS: 83
6 SOFTWARE: PatentIn Ver. 2.1
7 SEQ LENGTH: 2669
8 ORGANISM: Homo sapiens
9 TYPE: DNA
10 US-09-525-978B-82

Alignment Scores:
Seq. No.: 2,110-219 Length: 2669
Score: 2163.00 Mismatches: 294
Percent Similarity: 82.57% Conservative: 75
Best Local Similarity: 69.37% Mismatches: 85
Query Match: 14 Indels: 14
Gaps: 1

US-09-943-075A-2 (1-572) x US-09-525-978B-82 (1-2669)

Qy 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20
Db 92 ATGGAATGCTCTACTATTTCTCGGGATTTCTGCTCGCTGCAAGATTGGCACTTGAT 151
Qy 21 AlaAlaValArgPheArgAspValLeuGlyHisIsgLcInTyProAspHisMetArgGlu 40
Db 152 GCGGCGAACACGTTTCATGATGCTGTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 211
Qy 41 AsnAsnGlnLeuArgGlyTyPserSerAspGluAsnGluTrpAspGluGlnLeuTyPro 60
Db 212 CACATCAATTAATTAAGTGGTGTCTGTGATGAATAATGACTGAATGAATAATGACTGCTACCCA 271
Qy 61 ValTrpArgAspGlyHisArgTrpValPheSerTrpGluGlnIsgLcValArgGlnAla 80
Db 272 GTGTGGAAGGGGAGACATGAGCTGGGAATACTCTGGGAAGGGAGCGCTGTGAGCGG 331
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 332 GTCCGACGACCTGACTACACGCCCTCTGGGCTCCAAATATACATTTGGCGTGACCTG 391
Qy 101 ValPheProArgCysGlnTyGluAspAlaAsnGlyAsnIleValTyGluArgAsnCys 120
Db 392 ATATTCCCTAGATGCGCAAGGAAGATGCCATGTGCAACATAGTCTATGAGAAGAACTGC 451
Qy 121 ArgSerAspLeuGluAlaSerAspProTyValTyAsnTrpThrThrGlyAlaAsp 140
Db 452 AGAATATGGCTGTTATCTGCTGATCCATATGTTTAACTGAGACGATGCTGACAG 511
Qy 141 AspGluAspTrpGluAspAsnThrSerGlnGlnHisLeuArgPheProAspGlyLys 160
Db 512 GACATGAGGGGAATAATGGCAACGCCCAAGCCATCACTACGCTCTCCCTGATGGGANA 571
Qy 161 ProPheProArgProAlaIsgLcArgLysTrpGluPheValPheThrLeu 180
Db 572 CTTTTCTCACACCCCGGAGAGAGATGGAATTTCTACTAGCTCTTCCACACATT 631
Qy 181 GlyGlnTyPheGlnTyLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 632 GCGACATTTTCAGAAATATGGGACATGTGTACGTGTGAGAGATTTCTGTGACACAGCCAT 691
Qy 201 LeuThrValCysProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTy 220
Db 692 GTGACACTTGGGGCTCACTCATGGAAGTGAAGTCTGTACAGACATGACGCGGCATAT 751
Qy 221 IleProIleSerValValAspValTyValIleThrAspGlnIleProIlePheVal 240
Db 752 GTTCCCAACGACAGATGAAATGTACGTGTGTGACGATGACATGCTGTGTGTGTG 811
Qy 241 ThrMetCysGlnTyAsnAspArgAsnSerSerAspGluThrPheLeuAspLeuPro 260
Db 812 ACTATGTTCCGAAGACGATCGAAATTCATTCGACGAACACCTTCTTCAAAGATCTCC 871

Qy 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTySerAlaIleSer 280
Db 872 ATATGTTTGTATGCTTCATGATCATGATCTAGCCACTTCTCAATTATTTACCAATTAC 931
Qy 281 TyTrpAspPheGlyAspAsnThrGlyLeuPheValSerAsnAsnIleThrLeuAsn 300
Db 932 TACAGTGGAGCTTCGGGAGATATACATGAGCTGTGTTGTTCCACCAATCATCTGTGANT 991
Qy 301 HisThrTyValIleuAsnGlyPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Db 992 CACAGCATCTGCTCAATGGACCTTCAGCTTACCTGTAACCTGTAACCTGTAACCTGTAAC 1051
Qy 321 GlyProCysAspSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340
Db 1052 GGA-TGTTGTCGCCACGCCACGCCACGCCACGCCACGCCACGCCACGCCACGCCACGCC 1087
Qy 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyTrpSer 360
Db 1088 -----TCNMAACCCACCCCTCTTTAGGACTGCTGTGTGACACCC 1129
Qy 361 MetGluLeuSerAspIleSerAsnGluAsnGlyArgIleAsnArgTyGlyTyPheArg 380
Db 1130 CTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
Qy 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
Db 1190 GCAACATACATATTTAGAGGATCTTAGAGGTTTACATCTCCAGATGACAGATCTC 1249
Qy 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
Db 1250 CTGATGCGCGTCCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1309
Qy 421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
Db 1310 AGCATTCGACAGAGCTGTGTACCATCTTCTGACCCACCTGCGAGATCACCACGAC 1369
Qy 441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgArgAlaPhe 460
Db 1370 ACAGTCTCGACGCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1429
Qy 461 AsnGlySerGlyThrTyCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
Db 1430 AATGGCTCTGGAGCTACTGTGTGAACCTCACCTGGGAGTACACCAAGCTGGCTCTC 1489
Qy 481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
Db 1490 AGGACACCTGTTTGTCTCTGACAGACCCAGCTGCTGCTTTAGATGGCAAG 1549
Qy 501 GlyValLeuSerIleThrCysLeuAlaMetPheValThrIleLeuLeu 520
Db 1550 AGTGCCTGATCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGG 1609
Qy 521 TyTrpLysHisArgTyTrpTyLysProIleGlyAsnCysThrArgAsnValValLysGly 540
Db 1610 TACAATAACACAGGATACACACCATAGATAATAGTCTGGGAATGTGTCGACAGAG 1669
Qy 541 LysAlaLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
Db 1670 AAGAGCTGATGCTTCTTCAACCTGCGAAMACCGGTGTTCTTCCGGGGAACCGGAA 1729
Qy 561 LysAspProLeuLeuGlnAspLys 568
Db 1730 AAGATCGCTACTCAAAAACCAA 1753

RESULT 5

US-10-227-984-41
; Application US/10227984
; Sequence #1, NO. US2003002798A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary

1 APPLICANT: Goddard, Audrey
 2 APPLICANT: Godowski, Paul J.
 3 APPLICANT: Grimaldi, J. Christopher
 4 APPLICANT: Gurrey, Austin L.
 5 APPLICANT: Haddad, John
 6 APPLICANT: Stephan, Jean-Philippe F.
 7 APPLICANT: Natansbe, Colin L.
 8 APPLICANT: Wood, William I.
 9
 10 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 11
 12 TITLE REFERENCE: P350P1C79 US/10/227,884
 13
 14 CURRENT FILING DATE: 2002-08-26
 15
 16 PRIOR APPLICATION NUMBER: 10/119,480
 17
 18 PRIOR FILING DATE: 2002-04-09
 19
 20 PRIOR APPLICATION NUMBER: 60/059113
 21
 22 PRIOR FILING DATE: 1997-09-17
 23
 24 PRIOR APPLICATION NUMBER: 61/062287
 25
 26 PRIOR FILING DATE: 1998-03-20
 27
 28 PRIOR APPLICATION NUMBER: 60/063149
 29
 30 PRIOR FILING DATE: 1997-10-28
 31
 32 PRIOR APPLICATION NUMBER: 60/084103
 33
 34 PRIOR FILING DATE: 1997-10-31
 35
 36 PRIOR APPLICATION NUMBER: 60/069873
 37
 38 PRIOR FILING DATE: 1997-09-17
 39
 40 PRIOR APPLICATION NUMBER: 60/078910
 41
 42 PRIOR FILING DATE: 1998-03-20
 43
 44 PRIOR APPLICATION NUMBER: 60/079294
 45
 46 PRIOR FILING DATE: 1998-03-25
 47
 48 PRIOR APPLICATION NUMBER: 60/079656
 49
 50 PRIOR FILING DATE: 1998-03-26
 51
 52 PRIOR APPLICATION NUMBER: 60/079728
 53
 54 PRIOR FILING DATE: 1998-04-02
 55
 56 PRIOR APPLICATION NUMBER: 60/081819
 57
 58 PRIOR FILING DATE: 1998-04-15
 59
 60 PRIOR APPLICATION NUMBER: 60/081955
 61
 62 PRIOR FILING DATE: 1998-04-15
 63
 64 PRIOR APPLICATION NUMBER: 60/082804
 65
 66 PRIOR FILING DATE: 1998-04-15
 67
 68 PRIOR APPLICATION NUMBER: 60/084441
 69
 70 PRIOR FILING DATE: 1998-05-06
 71
 72 PRIOR APPLICATION NUMBER: 60/085323
 73
 74 PRIOR FILING DATE: 1998-05-13
 75
 76 PRIOR APPLICATION NUMBER: 60/085579
 77
 78 PRIOR FILING DATE: 1998-05-15
 79
 80 PRIOR APPLICATION NUMBER: 60/086392
 81
 82 PRIOR FILING DATE: 1998-06-02
 83
 84 PRIOR APPLICATION NUMBER: 60/089532
 85
 86 PRIOR FILING DATE: 1998-06-17
 87
 88 PRIOR APPLICATION NUMBER: 60/089538
 89
 90 PRIOR FILING DATE: 1998-06-17
 91
 92 PRIOR APPLICATION NUMBER: 60/089905
 93
 94 PRIOR FILING DATE: 1998-06-18
 95
 96 PRIOR APPLICATION NUMBER: 60/090472
 97
 98 PRIOR FILING DATE: 1998-06-24
 99
 100 PRIOR APPLICATION NUMBER: 60/090557
 101
 102 PRIOR FILING DATE: 1998-06-24
 103
 104 PRIOR APPLICATION NUMBER: 60/090691
 105
 106 PRIOR FILING DATE: 1998-06-25
 107
 108 PRIOR APPLICATION NUMBER: 60/090695
 109
 110 PRIOR FILING DATE: 1998-06-25
 111
 112 PRIOR APPLICATION NUMBER: 60/091982
 113
 114 PRIOR FILING DATE: 1998-07-07
 115
 116 PRIOR APPLICATION NUMBER: 60/095302
 117
 118 PRIOR FILING DATE: 1998-08-04
 119
 120 PRIOR APPLICATION NUMBER: 60/095318
 121
 122 PRIOR FILING DATE: 1998-08-04
 123
 124 PRIOR APPLICATION NUMBER: 60/095916
 125
 126 PRIOR FILING DATE: 1998-08-10
 127
 128 PRIOR APPLICATION NUMBER: 60/096146
 129
 130 PRIOR FILING DATE: 1998-08-11
 131
 132 PRIOR APPLICATION NUMBER: 60/096791
 133
 134 PRIOR FILING DATE: 1998-08-17
 135
 136 PRIOR APPLICATION NUMBER: 60/097986
 137
 138 PRIOR FILING DATE: 1998-08-26
 139
 140 PRIOR APPLICATION NUMBER: 60/098544
 141
 142 PRIOR FILING DATE: 1998-08-31
 143
 144 PRIOR APPLICATION NUMBER: 60/099596
 145
 146 PRIOR FILING DATE: 1998-09-09
 147
 148 PRIOR APPLICATION NUMBER: 60/099598
 149
 150 PRIOR FILING DATE: 1998-09-09
 151
 152 PRIOR APPLICATION NUMBER: 60/099803
 153
 154 PRIOR FILING DATE: 1998-09-10
 155
 156 PRIOR APPLICATION NUMBER: 60/099811
 157
 158 PRIOR FILING DATE: 1998-09-10
 159
 160 PRIOR APPLICATION NUMBER: 60/099812
 161
 162 PRIOR FILING DATE: 1998-09-10
 163
 164 PRIOR APPLICATION NUMBER: 60/099816
 165
 166 PRIOR FILING DATE: 1998-09-10
 167
 168 PRIOR APPLICATION NUMBER: 60/100038
 169
 170 PRIOR FILING DATE: 1998-09-11
 171
 172 PRIOR APPLICATION NUMBER: 60/100385
 173
 174 PRIOR FILING DATE: 1998-09-15
 175
 176 PRIOR APPLICATION NUMBER: 60/100390
 177
 178 PRIOR FILING DATE: 1998-09-15
 179
 180 PRIOR APPLICATION NUMBER: 60/100627
 181
 182 PRIOR FILING DATE: 1998-09-16
 183
 184 PRIOR APPLICATION NUMBER: 60/100848
 185
 186 PRIOR FILING DATE: 1998-09-18
 187
 188 PRIOR APPLICATION NUMBER: 60/100919
 189
 190 PRIOR FILING DATE: 1998-09-17
 191
 192 PRIOR APPLICATION NUMBER: 60/101477
 193
 194 PRIOR FILING DATE: 1998-09-23
 195
 196 PRIOR APPLICATION NUMBER: 60/101738
 197
 198 PRIOR FILING DATE: 1998-09-24
 199
 200 PRIOR APPLICATION NUMBER: 60/101741
 201
 202 PRIOR FILING DATE: 1998-09-24
 203
 204 PRIOR APPLICATION NUMBER: 60/101786
 205
 206 PRIOR FILING DATE: 1998-09-25
 207
 208 PRIOR APPLICATION NUMBER: 60/101916
 209
 210 PRIOR FILING DATE: 1998-09-24
 211
 212 PRIOR APPLICATION NUMBER: 60/101922
 213
 214 PRIOR FILING DATE: 1998-09-24
 215
 216 PRIOR APPLICATION NUMBER: 60/106178
 217
 218 PRIOR FILING DATE: 1998-10-28
 219
 220 PRIOR APPLICATION NUMBER: 60/106248
 221
 222 PRIOR FILING DATE: 1998-10-29
 223
 224 PRIOR APPLICATION NUMBER: 60/106464
 225
 226 PRIOR FILING DATE: 1998-10-31
 227
 228 PRIOR APPLICATION NUMBER: 60/106905
 229
 230 PRIOR FILING DATE: 1998-11-03
 231
 232 PRIOR APPLICATION NUMBER: 60/108787
 233
 234 PRIOR FILING DATE: 1998-11-17
 235
 236 PRIOR APPLICATION NUMBER: 60/108801
 237
 238 PRIOR FILING DATE: 1998-11-17
 239
 240 PRIOR APPLICATION NUMBER: 60/108849
 241
 242 PRIOR FILING DATE: 1998-11-18
 243
 244 PRIOR APPLICATION NUMBER: 60/112422
 245
 246 PRIOR FILING DATE: 1998-12-15
 247
 248 PRIOR APPLICATION NUMBER: 60/113296
 249
 250 PRIOR FILING DATE: 1998-12-22
 251
 252 PRIOR APPLICATION NUMBER: 60/113605
 253
 254 PRIOR FILING DATE: 1998-12-23
 255
 256 PRIOR APPLICATION NUMBER: 60/113621
 257
 258 PRIOR FILING DATE: 1998-12-23
 259
 260 PRIOR APPLICATION NUMBER: 60/115558
 261
 262 PRIOR FILING DATE: 1999-01-12
 263
 264 PRIOR APPLICATION NUMBER: 60/115565
 265
 266 PRIOR FILING DATE: 1999-01-12
 267
 268 PRIOR APPLICATION NUMBER: 60/115733
 269
 270 PRIOR FILING DATE: 1999-01-12
 271
 272 PRIOR APPLICATION NUMBER: 60/119549
 273
 274 PRIOR FILING DATE: 1999-02-10
 275
 276 PRIOR APPLICATION NUMBER: 60/123618
 277
 278 PRIOR FILING DATE: 1999-03-10
 279
 280 PRIOR APPLICATION NUMBER: 60/125259

2 PRIOR FILING DATE: 1999-03-19
 2 PRIOR APPLICATION NUMBER: 60/125775
 2 PRIOR FILING DATE: 1999-03-25
 2 PRIOR APPLICATION NUMBER: 60/126773
 2 PRIOR FILING DATE: 1999-03-29
 2 PRIOR APPLICATION NUMBER: 60/127887
 2 PRIOR FILING DATE: 1999-04-05
 2 PRIOR APPLICATION NUMBER: 60/130232
 2 PRIOR FILING DATE: 1999-04-26
 2 PRIOR APPLICATION NUMBER: 60/131370
 2 PRIOR FILING DATE: 1999-04-27
 2 PRIOR APPLICATION NUMBER: 60/131291
 2 PRIOR FILING DATE: 1999-04-28
 2 PRIOR APPLICATION NUMBER: 60/131445
 2 PRIOR FILING DATE: 1999-04-28
 2 PRIOR APPLICATION NUMBER: 60/134287
 2 PRIOR FILING DATE: 1999-05-14
 2 PRIOR APPLICATION NUMBER: 60/140650
 2 PRIOR FILING DATE: 1999-06-01
 2 PRIOR APPLICATION NUMBER: 60/140723
 2 PRIOR FILING DATE: 1999-06-22
 2 PRIOR APPLICATION NUMBER: 60/141037
 2 PRIOR FILING DATE: 1999-06-23
 2 PRIOR APPLICATION NUMBER: 60/144758
 2 PRIOR FILING DATE: 1999-07-20
 2 PRIOR APPLICATION NUMBER: 60/145698
 2 PRIOR FILING DATE: 1999-07-26
 2 PRIOR APPLICATION NUMBER: 60/146222
 2 PRIOR FILING DATE: 1999-07-28
 2 PRIOR APPLICATION NUMBER: 60/146963
 2 PRIOR FILING DATE: 1999-08-03
 2 PRIOR APPLICATION NUMBER: 60/149320
 2 PRIOR FILING DATE: 1999-08-31
 2 PRIOR APPLICATION NUMBER: 60/149638
 2 PRIOR FILING DATE: 1999-08-17
 2 PRIOR APPLICATION NUMBER: 60/151733
 2 PRIOR FILING DATE: 1999-08-31
 2 PRIOR APPLICATION NUMBER: 60/156418
 2 PRIOR FILING DATE: 1999-11-09
 2 PRIOR APPLICATION NUMBER: 60/166361
 2 PRIOR FILING DATE: 1999-11-16
 2 PRIOR APPLICATION NUMBER: 60/169445
 2 PRIOR FILING DATE: 1999-12-07
 2 PRIOR APPLICATION NUMBER: 60/169495
 2 PRIOR FILING DATE: 1999-12-22
 2 PRIOR APPLICATION NUMBER: 60/169835

Alignment scores:

Pred. No.: 2,71e-219
 Length: 2683
 Scores: 2162.00
 Best Local Similarity: 69.11%
 Best Global Similarity: 69.11%
 Query Match: 12
 Gaps: 2

US-09-943-075a-2 (1-372) x US-10-227-884-41 (1-2683)

Db 262 GTJTGGAAGCGGGAGACATGAGGTGGAAAACTCTGGAAGGAGGCGCTGTGCGAGCG 321
 Qy 81 ALIAIEthrSerAspSerProAlaLeuValGlySerAsnIleThrPheValAlaLeu 100
 Db 322 GTTCTGACGAGTACTACACGCGCTGTGGGCTCAATATACATTTGGCGTGAACCTG 381
 Qy 101 ValPheProAGTGCYSLnLysGluAspAlaAsodgLYasNIleValTYrGluArgAspCys 120
 Db 382 ATATCTCCCTAGATGCCNAAGAGAGATGCCCAATGGCAGCAATAGCTATGATGAGAACTGC 441
 Qy 121 ActSerAspLeuGluLeuAlaSerAspProTYrValTYrAsnTPhrGlyAlaasp 140
 Db 442 AGIAATGAGGCTGGTATTCTGCTGCTGATGTGTACAACTGGACAGCATGTGCAGAG 501
 Qy 141 AspGluAspTPGluAspAsnThrSerdInGlyInHISuArgPheProAspGlyIy 160
 Db 502 GACAGTGACGGGGAANAATGSCACGGCCAGAACCATCTAAGCTCTTCCTGTATGGGANA 561
 Qy 161 ProPheProArgProHISgLYArgLysLysTrpAsnPheValTYrValPheValThrLeu 180
 Db 562 CHTTTTCTCCACCCCGGAGAGAGATGGAATTTTCATCTACGCTCTCCACACACTT 621
 Qy 181 GLYGlnTYrPheGlnLysLeuGlyInCysSerAlaAaGValSerLISerLISerValasn 200
 Db 622 GSTCAGTATTTCAGAAATATGGCAGATGTTCAGTGAGAGTTCTGTGACACACCACT 681
 Qy 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHISgLYArgAlaTYr 220
 Db 682 GTGACATTTGGGCTCTACATCTAGAGATGACTGTCTACAGAGNCAATGGAGGCAAT 741
 Qy 221 IIEP-oIleSerLysValLysAspValTYrValIleThrAspGlnIleProIlePheVal 240
 Db 742 GTTCTCCATCGCACAGTGAAGATGTGTACGTGGTAACAGATCAGATCTCTGTGTGTG 801
 Qy 241 ThrMetTYrGlnLysAspArgPheSerAspGlnIleThrPheLeuArgAspLeuPro 260
 Db 802 ACTATGTTTCAGAGACATGCAATTCATCCGAGAAACCTCTCTCAAGATCTCCG 861
 Qy 261 IIEPhePheAspValLeuIleHISAspProSerHISpHeLeuAsnTYrSerAlaIleSer 280
 Db 862 ATATGTTTGTGCTGTGATCTGATCTACCTCTCTCTCATATTCTACCATTAAC 921
 Qy 281 TYrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHISThrLeuAsn 300
 Db 922 TACNAGTGGAGCTGGGGATATATCTGCTGTTTGTTCACCAATCATCTGTGAT 981
 Qy 301 HIsThrTYrValIleAsnGlyThrPheAsnPheAsnLeuThyValGlnThrAlaValPro 320
 Db 982 CAGAGTATCTCTCAATGGACCTTCACGCTTCACTCACTGTAAGCTGCAGACCA 1041
 Qy 321 GlyProCysProSerProThrProSerProSerSerSerSerProSerProAlaSer 340
 Db 1042 GACCTGTCTCCGCGCCGCCGCCACCCACCCAGA-----CCTTCA 1080
 Qy 341 SerProSerProSerLeuSerThr-----ProSerProSerLeuMetPro 355
 Db 1081 AAACCCACCCCTCTTTAGCACTACTATAAATCTTATGATTCAACCCAGGAGCT 1140
 Qy 356 ThrGlyTYrLysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg 375
 Db 1141 ACTGTGACAAACCCCGGAGCTGAGTAGGATCTCTGTAGTAACATGCCAGTTAACAGA 1200
 Qy 376 TYrGlyTYrPheArgAlaThrIleValAspGlyIleLeuGluValAsnIleIle 395
 Db 1201 TATGCGCCATTAAGCCACCATCAATGTAGAGGAATCTTAGAGTTTAACATCTC 1260
 Qy 396 GluValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
 Db 1261 CAGATGACAGAGCTCTGTGATGGCGGCTGCAAGCTGCTTAATAGACTTTGTC 1320
 Qy 416 ValThrCysLysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
 Db 1321 GTGATCTGCCAAGGAGGACATCTCCACGAGAGTCTGTACCATCTCTTGACCCACCTGC 1380


```

; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-05
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-01
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-04
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-26
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638

```

```

; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Alignment Scores:
      No.:      2,71e-219      Length:      2683
      Score:      2162.00      Matches:      398
      Percent Similarity:      82.72%      Conservative:      78
      Best Local Similarity:      69.11%      Mismatches:      87
      Query Watch:      70.63%      Indels:      12
      DB:      9      Gaps:      2

US-09-943-075a-2 (1-572) x US-10-230-163-41 (1-2683)

Qy 1 MetGluserLeuCysGlyValLeuValPheLeuLeuAlaIaGlyLeuProLeuIn 20
Db 82 ATGGAATCTCTACTATTTCTCTGGGATTTCTCTGGCTGCAAGATTCGCACCTGAT 141
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisIcLugInTyRProAspHisMetArgGlu 40
Db 142 GCCGCCAAGCATTTTCATGTCGTCTGGGCATGCAAGACCTTCTGCTTACATGAGGAG 201
Qy 41 AaaAsnGlnLeuArgGlyTyRSerSerAspGluAsnGluTyRPaSpGluGlnLeuTyRPro 60
Db 202 CACMATCAATTAATGGCTGCTCTCTGATGAATAAGTACGGATGAATAAGTACGCCCA 261
Qy 61 ValTyRPaRgArgGlyGluGlyValGtyRtyLysAspSerTyRGLuGlyValArgValGlnAla 80
Db 262 GTGTCGAAGCGGGAGACATGAGGTGGGAAAACCTCTGGAGGAGCGGTGTGCGAGCGC 321
Qy 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValIvalAsnLeu 100
Db 322 GTCTCTGACAGTGACTACCAAGCCTCTGGGGCTCTCATATATACATTTGCGGTGAACCTG 381
Qy 101 ValPheProArgCysGlnLysGlnAspAlaAsnGlyAsnIleValTyRGIuArgAsnCys 120
Db 382 ATATTCCTCTAGATGCCAAGAAGAGATGCCAATGGCAACATAGTCTATGAGAAGACATGC 441
Qy 121 ArgSerAspLeuGluLeuAlaAspArgProTyRValTyRAsnTyRThrThrGlyAlaAsp 140
Db 442 AGAATAGAGCGTGGTGTATCTGCTGCTATCCGTATGTTTACACTGCAACATGGTGCAGAG 501
Qy 141 AspGluAspTyRProGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
Db 502 GACAGTGACGGGAAATATGGCCCGGCCAAGCCATACACCTCTCCCTGATGGGAAA 561
Qy 161 ProPheProArgProHisGlyValGlyLysTyRTPasPheValTyRValPheIleThrLeu 180
Db 562 CCTTTCTCCACCACCCCGGATGGAGAAGATGGAATTCATCTACGTCCTCCACACATT 621
Qy 181 GlyGlnTyRPhesGlnTyLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
Db 622 GGTCAGTATTTCCAAATATGGGAGCATGTCCTGCGARGTTTCTGACACACCCCAAT 681
Qy 201 LeuThrValGlyProGlnValLelGluValIleValPheArgHisGlyArgAlaTyR 220
Db 682 GTGACACTTGGCGCTCAACTCATGGAGTGTGCTGTCTACAGAAGACATGGAGGCGCATAT 741
Qy 221 IleProIleSerLysValLysAspValTyRValIleThrAspGlnIleProIlePheVal 240
Db 742 GTTCCATCCGACAGTGAAGAAATGTAGTGGTGTACNGATTCATGTTCTGTGTGTG 801
Qy 241 ThrMetTyRGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260

```

```
Db 802 ACTATGTTCCAGAGACGATCGAAATTCATCCGACGAAACCTTCTCCAAAGATCTCCCC 861
Qy 261 ILePhePheAspValLeuIleHisAspProSerHisPheLeuAanTySerAlaIleSer 280
Db 862 ATTATGTTGATGTCCTGATTCATGATCAATCACTAGCCACTCTCTCAATATATCTACCAATAC 921
Qy 281 TydTySerAspPheGlyPheAspThrGlyLeuPheValSerAsnAsnHisLeuLeuAsn 300
Db 922 TACAAGTGAGCTTCGGGATAACTAGCTGCTCTGTGTGTTGTCACCAATCATATCTGTGAAT 981
Qy 301 HisTrpTrpValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Db 982 CACAGCTATGTCCTCAATGACACCTTCAGCCTCTACTCTACTGTGAAGAGTCGACGCCA 1041
Qy 321 GlyProCysAspProThrProSerProSerSerSerThrProSerProAlaSer 340
Db 1042 GSACCTTGTCGCCACGCCACCCACCCACCCAG-----CCTTCA 1080
Qy 341 SerProSerProThrLeuSerThr-----ProSerProSerLeuMetPro 355
Db 1081 AATCCACCCCTTTTATGACACTCTCTMAATCTATGATCAACACCCAGGACCT 1140
Qy 356 ThrGlyTyTrpSerMetGluLeuSerAspIleSerAsnGluLeuAsnCysAspGlnAsnTrp 375
Db 1141 ACTGTGTGACACCCCTCGAGCTGAGTAGGATCTCTGCTGAAACTGCCAGATTAACAGA 1200
Qy 376 TygGlyTrpThrArgAlaThrThrIleValAspGlyIleLeuGluValIleHisLe 395
Db 1201 TATGSCACTTTCACCCACCACATCATCTGATAGAGGAATCTTAGGTTTACATCATC 1260
Qy 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
Db 1261 CAGATGACAGACCTCTGATGCGCGGTGCCATGCGCTGAAAGCTCCCTAATGACTTTGTC 1320
Qy 416 ValThrCysValysGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
Db 1321 GTGACTGTCGACAGGAGGATCTCCACCGAGTCTGTACCATCTTCTGACCCACCTGC 1380
Qy 436 GlnIleAlaGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuSer 455
Db 1381 GAGTTCACCCAGAACAGTCTGACGCCCTCTGATGTGATGAGATGTGTCTGCTGCT 1440
Qy 456 ValArgAlaGlnAlaPheAsnGlySerGlyThrTyTrpCysValAsnPheThrLeuGlyAsp 475
Db 1441 GTGAGAGAACCTTCAATAGTGGTCTGGGAGTACTGTGTGAACCTCACCTCGGGGATGAC 1500
Qy 476 AlaSerLeuAlaLeuThrSerAlaAlaIleSerIleProGlyIysAspLeuGlySerPro 495
Db 1501 ACAAGCTGGCTCTCAGCAGCAGCCTGTTCTGCTGACAGACCCAGCCGCTGCT 1560
Qy 496 LeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMet 515
Db 1561 TTAAAGATGACCAACAGTGCCTGATCTCCGTGTGCTGCTTGGCCATATTTGTCATGTG 1620
Qy 516 ValThrIleLeuLeuTyTrpIysValHisIysThrTyIysProIleGlyAsnGlyThr 535
Db 1621 ATCTCCCTCTGTGTGTACAAAAACACAGGAATACAAACCAACCAATAGTCTCGG 1680
Qy 536 AsnValValIysGlyIysGlyLeuSerValPheLeuSerHisAlaIysAlaProPheSer 555
Db 1681 ATGTGTGTCTGACGAAGAGCCCTGAGTGCTTCTTCAACCGTCGCAAGAGCGGTCTTTC 1740
Qy 556 ArgGlyAspArgGlyAspProLeuLeuIleGluAspIys 568
Db 1741 CCGGGAACACGAGGAAGGATCCGCTACTCAAAACCA 1779
```

RESULT 7

US-10-218-631-41

Application US/10218631

Inventor: Baker, Kevin P.

Publication No. US20030045687A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Desnoyers, Luc

```
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Slach, Victor
APPLICANT: Stachurski, J. Philippe F.
APPLICANT: Walenabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530P1G14
CURRENT APPLICATION NUMBER: US/10/218,631
PRIORITY NUMBER: 08-12
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 246
SEQ ID NO 41
LENGTH: 2683
TYPE: DNA
ORIGIN: Homo Sapien
US-10-218-631-41
Alignment Scores:
Score: 2,71e-219 Length: 2683
Pred. No.: 2162.00 Matches: 596
Percent Similarity: 82.72% Conservative: 9
Gap Penalty: 10 Indels: 12
Gaps: 2
DB: 9
```

US-09-943-075A-2 (1-572) x US-10-218-631-41 (1-2683)

```
Qy 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuAlaIleAclyLeuProLeuGln 20
Db 82 ATGCAATGCTCTACTATTTCCTGGGATTTCTGCTCTCGCTGCTGAGATTCGCACTTGAT 141
Qy 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyProAspHisMetArgGlu 40
Db 142 GCCGCCCAACGATTTTCATGATGCTCTGGGCATGAAGACCTTCTGCTTACATAGGGAG 201
Qy 41 AsnSerGlnLeuArgGlyTrpSerSerAsnGluAsnGluTyPheGluGlnLeuTyPro 60
Db 202 CACAATCAATTAATGGCTGCTCTCTGATGAAATGACTGGGAATGAATAAATCTCTACCA 261
Qy 61 ValTrpArgArgGlyGluGlyTrpIysAspSerTrpGluGlyArgValGlnAla 80
Db 262 GTGTGGAGCCGGGAGCATGATGGTGGTGAATAAATCTCTGGAAGGAGGCGCTGTGAGGG 321
Qy 81 AlaLeuThrSerAsnSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
Db 322 GTCTTGACCTGATCTCACACGCCCTCTGGGCTCAAAATAAATAACATTTTCGGTGAACCTG 381
Qy 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyGluArgAsnCys 120
```


DB	382	ATATTCCCTAGATGCCAAGGAGATGCCATGCCACATATGTCTATGAGAGAACTGC	441
OY	121	ArgSerAspLeuGluLeuAlaSerAspGlyValTyrAsnTrpThrGlyAlaAsp	140
DB	442	AGAAATGAGCGTGTATCTGCTGATGATCTGATGTTTACAACTGACAGCATGCTCAG	501
OY	141	AspGluAspTrpCllAspAsnThrSerGlnGlyHisLeuArgPheProAspGlyLys	160
DB	502	GACAGTGACGGGGAATAGCCACGGCCCAAGCCATCATACGCTTCTCCGTGATGGGAA	561
OY	161	ProPheProProHisLysGlyArgLysTyrTrpAsnPheValTyrValPheLisThrLeu	180
DB	562	CCCTTTCTCTCCACGCCCGGATGAGAGAAAGTGAATCATCTAGCTGTCCACACAT	621
OY	181	GlyGlyTyrPheClnGlyLeuGlyCysLeuAlaSerLysAspThrValAsp	200
DB	622	GGTCAGTATTTCACAAATTGGCAGGATGTCTCAGTCAGAGGTCTCTGTGAACACAGCCAT	681
OY	201	LeuThrValGlyProGlnValIleValPheArgGHIserGlyArgLeuTyr	220
DB	682	GTGCACTTGTGGCGTCAACTCTGAGAGTACGTCTCTACGAAGCATATGACGGCATAT	741
OY	221	IleProTLeuSerLysValValAspGlyTyrValIleThrAsnGlnIleProLlePheVal	240
DB	742	GTTCCTCATGCCAAGTGAAGATGTGTACGTGTACACAGTACAGTCCGTGTGTGTG	801
OY	241	ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro	260
DB	802	ACTAGTGTCCAGAACACCATCGAATTCATCCGACGAACATCTCTCAAGATATCTCCG	861
OY	261	LlePhePheAspAlaLeuIleHisAspProSerHisGlyLeuLeuAspThrSerAlaL	280
DB	862	ATTATGTGTGATCTCTGATCATGATCTAGCCACTCTCCATTAATTTCTACCATTAAC	921
OY	281	TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn	300
DB	922	TACAGTGGACGCTCGGGGATATACCTGGCGTGTGTGTCTCCACCACATCATCTGAT	981
OY	301	HisThrTyrValLeuAsnGlyThrPheAsnPheAsnGluThrValGlnThrAlaVal	320
DB	982	CACAGCTATGTCTCATGSAACCTTCAGCCTTAACCTCACTGTGAAGCTCGACACCA	1041
OY	321	GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer	340
DB	1042	GGACTGTGTGTGGTCACCGCCACCCACACACACACACACACACACACACACACAC	1080
OY	341	ProSerProSerProSerProSerProSerProSerProSerProSerProSerProSer	355
DB	1081	AAACCCACCCCTCTCTTTAGCACTACTCTAAAATCTTATGATTCAAAACACCCAGGACCT	1140
OY	356	ThrGlyTyrLysSerMetGluLeuSerAspLleSerAsnGluAsnCysArgLleAsnArg	375
DB	1141	ACTGTGGACACACCCCTGGAGTGTAGTACAGTCTGTGAGAACTCCGATTAACAGA	1200
OY	376	TyrGlyTyrPheAlaThrThrLemhrLleValAspGlyLleLeuGluValAsnLleIle	395
DB	1201	TATGGCCACTTTCACAGCCACCATCAATATTGTAGGGGATCTTATAGAGTTCACATCAT	1260
OY	396	GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle	415
DB	1261	CAGATCGACAGCCTCTGATCGCGCCGACGCGCCGACGCGCGCGCGCGCGCGCGCG	1320
OY	416	ValThrCysGlyAsnThrProThrClnAlaCysThrLisLleSerAspThrCys	435
DB	1321	GTGACCTGCCAAGGGAGCATTCGCCAGGAGGTCTGTACCATCATTTCTGACCCGACCTGC	1380
OY	436	GlnTleIleGlnAsnArgValCysSerProValAlaValAspGluLeuCysLeuLeuSer	455
DB	1381	GAGATACCCGAACACATGCGACGCCCTGTGGATGTGGATGATGTGTGTGTGTGTGT	1440
OY	456	ValArgGlnAlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyLysAsp	475

Score:	-2162.00	Matches:	396
Percent Similarity:	82.72%	Conservative:	78
Best Local Similarity:	69.11%	Mismatches:	87
Query Match:	70.63%	Indels:	12
DBs:	9	Gaps:	2
US-09-943-075A-2 (1-572) x US-10-230-338-41 (1-2683)			
QY	1 MetGluSerLeuCyScylValLeuValPheValPheValLeuAlaLeuAlaLeuAlaLeuProLeuLeuLIn	20	
QY	82 MYGAAGTGTCTACTATTTCTGGGATTTCTGCTTGGCTGGCAAGATGCCACTTCAT	141	
QY	21 AlaAlaLysArgPheArgAspValIleuGlyIleSGuInuYrProAspHisMetArgLys	40	
QY	142 GCGCCCAACAGATTTCTATGATCTGTGGCAATGAAGAACTTCGTCTACATGAGGAG	201	
QY	41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGlnTrpAspGluGlnLeuYrPro	60	
QY	202 CACAATCAATTAATGCGCTGGCTCTGTGTAAGAAATGACTGGATGAAGAAATCTACCCA	261	
QY	61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla	80	
QY	262 GTGTGAGT	321	
QY	81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnGlnLeuPheValHisAsnLeu	100	
QY	322 GTCTGTGACAGCTGACTCACCAGCCCTCTGGGCTCAATATACATATTCGGGTGAACCTG	381	
QY	101 ValPheProArgCysInLysGluAspAlaAsnGlyAsnGlyValYrGluArgAsnCys	120	
QY	382 ATATCTCTGT	441	
QY	121 ArgSerAspLeuGlnLeuAlaSerAspProYrValYrAsnTrpThrThrGlyAlaAsp	140	
QY	442 AGAATGAGGCGCTGTTATCTGCTGTATGCTGTATGTTTACAACCTGGACATCATGTGACAG	501	
QY	141 AspGluAspTrpGluAspAsnThrSerGlnGlyInLysIleuArgPheProAspGlyLys	160	
QY	502 GT	561	
QY	161 ProPheProArgProHisIleArgLysLysTrpAsnPheValYrValPheHisThrLeu	180	
QY	562 CTTTTCCTCTCACACCCCGGTGGAGAAGATGGAAATTCATCTACGCTCTCCACACCTT	621	
QY	181 GlyGlnYrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn	200	
QY	622 GTCATANTTCCAGAAATTTGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	681	
QY	201 LeuThrValGlyProGlnAlaMetGlnValIleValPheArgArgIleGlyArgLayr	220	
QY	682 GTGACATCTGGGCTCAACTCATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	741	
QY	221 IleProLysSerValLysAspValYrValIleThrAspGlnIleProIlePheVal	240	
QY	742 GTTCCCATGCGACACCAAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	801	
QY	241 ThrMetYrClnLysAsnAspArgAsnSerAspGlyTrpPheLeuArgAspLeuPro	260	
QY	802 ACTATGTTCAGAGAGCATGCAATATCATCGCAAGCAACCTCTCTCAAGATCTTCCC	861	
QY	261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTrpSerAlaIleSer	280	
QY	862 ATATGATTTGT	921	
QY	281 YrLysTrpAsnPheCylAspAsnThrGlyLeuPheValSerAsnHisIleThrLeuAsn	300	
QY	922 TACAAAGTGGAGCTTCGGGGATTAATCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGT	981	
QY	301 HisThrYrValLeuAsnGlyThrPheAsnPheAsnThrValGlnThrAlaValPro	320	
QY	982 Gln	1041	
QY	321 GlnProCysProSerProThrProSerProSerSerThrSerProSerProAlaIleSer	340	

```

: PRIOR FILLING DATE: 2002-04-09
: PRIOR APPLICATION NUMBER: 60/059113
: PRIOR FILLING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/062287
: PRIOR FILLING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/063549
: PRIOR FILLING DATE: 1997-10-26
: PRIOR APPLICATION NUMBER: 60/064103
: PRIOR FILLING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/069873
: PRIOR FILLING DATE: 1997-12-17
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILLING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILLING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILLING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILLING DATE: 1998-03-27
: Remaining Prior Application data removed - See file Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 246
: SOURCE: GenBank
: LENGTH: 2683
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-230-414-41
Alignment Scores:
Score: 2,716-219
Score: 2162.00
Matches: 395
Conservative: 78
Best Local Similarity: 82.72%
Mismatches: 87
Indels: 12
Query Match: 76.63%
Gaps: 9
US-09-943-075A-2 (1-572) x US-10-230-414-41 (1-2683)
QY 1 MetGluSerLeuGlyValLeuValPheLeuLeuAlaGlyLeuProLeuGln 20
DB 82 ATGAGATGCTCTACTACTATTTCTCTGGGATTTCTCTGCTGCGCAAGATTGCCACTTGA 141
QY 21 AlaAlaValSerPheAAspValLeuGlyHisAlaGlyLeuTyPheAspHisMetArgGlu 40
DB 142 GCGGCCAAACGATTTCATCATGCTGCTGGCAATGAAGACCTCTCTCTTACATGGGAG 201
QY 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluGlnGluTrpAspGluGlnLeuTyPro 60
DB 202 CACATCAATTAATAGCTGCTCTCTGTGTAAGATTAAGTCAATGCAATCACTCTACCA 261
QY 61 ValTrpArgArgGlyGlyGlyValGlyTrpGlyAspSerTrpGlyGlyValGlyAlaGln 80
DB 262 GTGTGGAAAGGGGGAGACATGAGGTGGAAATCTCTGGAGGAGGCGCTGTGAGGGG 321
QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
DB 322 CTCCTGACCATGTACTCACCAGCCTCTGCTGGCTCAATATTAACATTTGCGGTGAACCTG 381
QY 101 ValPheProGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 120
DB 381 ATATTCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 121 AsnSerAspLeuGluLeuAlaSerAspGlyValTyValTyAsnThrThrGlyAlaAsp 140
DB 442 AGAATAGAGCGCTGTTTATCTGCTGATCGCTATGTTTACACTGGACAGCATGCTGACAG 501
QY 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
DB 502 GACATGTCGGGGAATTAATGACCGCCCAACGCAATCACTACCTCTCTCTGATGGGAAA 561
QY 161 ProPheProArgProHisHisGlyValGlyLysTrpAsnPheValTyValPheHisThrLeu 180
DB 562 CATTTCCTCACCACCCCGGATGGAGAGATGAATTTCTACCTACGCTCTCCACACACTT 621

```

```

OY 181 GlyGlnTyPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
DB 622 GGTCACTATTTCAGAAATTTGGCAGCATGTTCCAGTCAGAGATTTCGTGTGACACACCAAT 681
OY 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyValArgAlaTy 220
DB 682 GTGACACTTGGGCTCACTCACTGAGAGTACTGCTCTACAGAGACATGAGCAGGCAAT 741
OY 221 IleProIleSerLysValLysAspValTyValIleLeuAspGlnIleProIlePheVal 240
DB 742 GTTCCCATCGCACAGTGAAGATGTGTACGTGTAAACAGATCAGATTCTGTGTGTGTS 801
OY 241 ThrMetTyRGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgPheLeuPro 260
DB 802 ACATGTTTCAGAGAAACCAATTAATTCATCCGAGCAACCTCTCTCAAGATCTCCCC 861
OY 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTySerAlaIleSer 280
DB 862 ATTATGTTTGATGCTGTGATTCATGATGCTGACCACTCTCTCAATATTATACCATTAAC 921
OY 281 TyIleTPAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
DB 922 TACATGAGCTCTGGGATTAATACGTGGCGCTGTTGTTTCACCAATCACTGTGAT 981
OY 301 HisThrTyValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnIleAlaValPro 320
DB 982 CACACCTGATGCTCAATAGGAACCTTCAGCCTTAACCTCACTGTGAACACTGCAGCACA 1041
OY 321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340
DB 1042 GCACCTTGTGCGCACCGCCGACCAACACCCAGCA 380
OY 341 SerProSerProThrLeuSerThr 355
DB 1081 AAAACCCCTCTCTTACCACTACTCTAAATCTTATGATTTCAACACCCAGGAGCT 1140
OY 356 ThrGlyTyLysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArg 375
DB 1141 ACTGTCGACACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGA 1200
OY 376 TyroIleTyPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIle 395
DB 1201 TATGSCCATTTGAAGCCACATCAATTTGTAGAGGGAATCTTAGAGTTAAATCATC 1260
OY 396 GlnValAlaAspValProIleProThrThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
DB 1261 CAGATGACAGCTCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1320
OY 416 ValThrCysGlyGlyAlaThrProHisGlyAlaValThrIleLeuSerAspProThrCys 435
DB 1321 GTGACCTGCCACAGGAGCAATTCGCCAGAGGCTGTGACCATCATTTCTGACCCACCTG 1380
OY 436 GlnIleAlaGlnArgValCysSerProValAlaValAlaAspGluLeuLeuSer 455
DB 1381 GAGATGACCCAGACAGCTGTGACAGCTGTGATGTGATGTGATGTGATGTGATGTGATGT 1440
OY 456 ValTrpGlnAlaPheAsnGlySerGlyThrTyTyValValAsnPheThrLeuAspArg 475
DB 1441 GTGAGGACACCTCAATGGTCTGGACGCTACTGTGTGAACCTCCACCTGGGGGATGAC 1500
OY 476 AlaSerLeuAlaThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerPro 495
DB 1501 ACAAGCTGCTCTACAGCAGCCCTGATTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
OY 496 LeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMet 515
DB 1561 TTAAGGTGGCAACAGTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
OY 516 ValThrIleLeuLeuTyLysValHisLysThrTyLysProGlyLeuGlyAsnCysThrArg 535
DB 1621 ATCTCCCTCTGCTGTACAAAAACACAAAGAAATACCAACCAATAGTAATGTCTGGG 1680
OY 536 AsnValValLysGlyLysGlyLeuSerValPheLeuSerHisAlaValAlaProPheSer 555

```


1519 TACAAACACACAGATACACACCATAGAAATAGCTCTGGGATGTGTGTCAGAAC 1578
 541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaLysPheSerArgGlyAspArgGlu 560
 1579 AAGAGCGCTGAGTGTCTTCTCAACCGTGCAGAAAGCGGTGTCTCCCGGAAACAGGAA 1638
 561 LysAspProLeuLeuGlnAspLys 568
 1639 AAGGATCGCTACTCAAAAACCA 1662

RESULT 11
 US-10-039-272-1 Application US/10039272
 ; Sequence 1, Application US/10039272
 ; General Information
 ; APPLICANT: RAMESHVAR, Penela
 ; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE
 ; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
 ; CURRENT APPLICATION NUMBER: US/10/039,272
 ; PRIOR APPLICATION DATES: 2001/03/20
 ; PRIOR APPLICATION NUMBERS: US 60/241,891
 ; PRIOR FILING DATE: 2000-10-20
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 2661
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens
 US-10-039-272-1

Alignment Scores:
 Pred. No.: 1,54e-216 Length: 2661
 Score: 2136.00 Matches: 393
 Best Local Similarity: 62.19% Conservative: 82
 Query Match: 69.78% Indels: 15
 DB: 9 Gaps: 1

US-09-943-075a-2 (1-572) x US-10-039-272-1 (1-2661)

QY 1 ValGluSerLeuGlyValValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
 DB 60 ATGGAAATGCTCTACTATTCTCGGGATTCTGCTGCTGCTGCAAGATTGCCATTGAT 119
 QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspLysMetArgGlu 40
 DB 120 GCGCCACACCAATATTCATGATGCTGGGCAATGAAGACCTCTCTCATGAGGAG 179
 QY 41 AsnAsnLeuSerAspGluAspGlyTyrPheSerArgGluAsnGlnTyrAspGluLeuTyrPro 60
 DB 180 CACATCAATATAATAGGCTGGCTCTCTGATGAATATGACTGGAATGAAATCTTACCCA 239
 QY 61 ValTyrPheArgGlyGlyArgTyrPheAspSerTyrPheGlyArgValGlnAla 80
 DB 240 GTGTGGAAAGCGGGAGACATGAGCTGGTGAATATCTCTCGAGAGCGCTGTGCGGCG 299
 QY 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnLeuThrPheValValAsnLeu 100
 DB 300 GTCTCGACCAAGTCAACAGCCCTGGGGCTCAAAATATACATTTCCGGTGAACCTG 359
 QY 101 ValPheProArgGlyGlyGlyValAsnAlaAsnGlyAsnTyrGluArgAspCys 120
 DB 360 ATATTCTAGATGCCAAAGGAATGCCAATGSCACATATCTATGAGAGAATCTC 419
 QY 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnThrThrGlyAlaAsp 140
 DB 420 AGAATAGAGCTGGTGTATCTGCTGATGCTATGTTTACATGTGACAGATGGTCAGAG 479
 QY 141 ArgGluAspTyrGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
 DB 480 GACAGTACCGGGAATATGACCGGCAAGCATCATCATCATGATCTCTTCCCTGATGGCAA 539

QY 161 ProPheProArgProHisHisGlyArgLysTyrPheAsnPheValTyrValPheHisThrLeu 180
 DB 540 CTTTTCCTTCACCCACCCGATGGAGAGATGATTCATCTAGCTTCTTCACACACT 599
 QY 181 GlyTyrPheThrPheLeuGlyGlnGlyCysSerAlaArgValSerLeuAsnThrValAsn 200
 DB 600 GGTCAATATTCCAGAAATTTGGAGAGATGTCAGTGAGAGTTCTGTGACACACAGCAAT 659
 QY 201 LeuThrValGlyProGluValValMetGlyValValValPheArgPheGlyArgValTyr 220
 DB 660 GTGACACTTGGCGCTCAACTCATGTGAAGTGTCTGTACGAAGACATGSCAGCGCATAT 719
 QY 221 IleProIleSerIleValValAspValTyrValIleThrAspGlnIleProIlePheVal 240
 DB 720 GTCCCATCGCAGAGATGAGATGTGACGTGTACACAGATCAATGTCTGTGTGTG 779
 QY 241 ThrMetTyrGlnLysAsnArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
 DB 780 ACTATGTCTCAGAGAACAATCAAAATTCATCCGAGAAACCTTCCC-AAAGATCTCCCC 838
 QY 261 IlePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
 DB 839 ATTATGTGTGATCTGATCATGATCATCTAGCACTCTCTCTCAATTTATCTACCATTAAC 898
 QY 281 TyrTyrTyrPheAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300
 DB 899 TACAAGTGGAGCTCGGGGATTAATCTGGCTGTGTGTTCACCAATCATCTACTGTGAAT 958
 QY 301 HisThrTyrValLeuSerGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
 DB 959 CACACGTTGTGCTCAATGGAACTTTACGCTTACCTCACTGTCAAACTGCAACGCCA 1018
 QY 321 GlyProCysProSerProThrProSerProSerSerThrSerProSerProAlaSer 340
 DB 1019 GACCTTGTGCGCCACCGCCACCCACCCACCGCT----- 1054
 QY 341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyTyrLysSer 360
 DB 1055 -----TCAAAACCCACCCCTCTTTAGGACCTGCTGTGACCAACCCC 1096
 QY 361 MetGluLeuSerAspLysLeuAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380
 DB 1097 CTGACGTAGTAGGATTCCTGATGANAATGCCAGATTACAGATATGCCACTTCAA 1156
 QY 381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleValAlaAspVal 400
 DB 1157 GGCACCAACAATTTAGAGAGGATTTAGAGGTTCATCATCCAGATCAGATCAGACATC 1216
 QY 401 ProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
 DB 1217 CTGATGCGCGTGCATGGCTGGAAGCTCCCTAATAGACTTTTGTGTGACCTGCCAAGG 1276
 QY 421 AlaThrProThrGluAlaCysThrIleIleSerAspProCysGlnIleAlaGlnAsn 440
 DB 1277 AGCATTCACCGAGTGTATACALASPGLuLeuGlyCysGAGCTGCGAGATCACCGAAC 1336
 QY 441 ArgTyrCysSerProValAlaValAspGluLeuCysLeuSerValArgArgAlaPhe 1360
 DB 1337 ACATCTTCACGCCCTCTGGATGTGGATGATGATGCTGCTGCTGCTGAGAGAACCTTC 1396
 QY 461 AsnGlySerGlyTyrTyrCysValIleAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
 DB 1397 AATGGCTGTGGAGCATGTGTGTGTGACCTCACTGGGGAGACACACACACCTGCTCTC 1456
 QY 481 ThrSerAlaLeuIleSerIleIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
 DB 1457 ACGAGACACCTGATTTCTGCTGTGACAGAGACCGCTCGCTGCTTAAAGATGGCCAAC 1516
 QY 501 GlyValLeuIleSerIleGlyCysLeuAlaSerPheValThrMetValThrIleLeuLeu 520
 DB 1517 AGTGCCTGATCTCCGTTGGCTGTGGCATATTTCCTCACTGTGATCTCCCTCTTGTGG 576
 QY 521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540

```

1577 TACAAAAACACAGGAATACACACCCATAGAAAATAGTCTCGGAGATGTGGTACAGAGC 1636
|||||
541 LysGlyLeuSerValPheLeuSerHisAlaValAlaProPheSerArgGlyAspArgGlu 560
|||||
1637 AANGGCTGAGTGTCTTCACAGCGTGCAGAAAGCGGTCTTCGCGGGAACCAAGAA 1696
|||||
561 LysAspProLeuLeuGlnAspLys 568
|||||
1697 AANGATCCGCTACTCAAAACCA 1720
|||||
RESULT 12
US-09-812-238B-1
: Sequence 1, Application US/09812238B
: Patent No. US20020169132A1
: GENERAL INFORMATION:
: APPLICANT: Nicolette, Charles
: TITLE OF INVENTION: THERAPEUTIC ANTI-MELANOMA COMPOUNDS
: CURRENT INVENTION NUMBER: US/09/812,238B
: CURRENT APPLICATION NUMBER: US/09/812,238B
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSP for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2130
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (22)...(2004)
US-09-812-238B-1
Alignment Scores:
Seq. No.: 1,56e-56 Length: 2130
Score: 627.50 Matches: 176
Percent Similarity: 41.85% Conservative: 96
Best Local Similarity: 27.08% Mismatches: 205
Query Match: 20.50% Indels: 173
DB: 9 Gaps: 19
US-09-943-075a-2 (1-572) x US-09-812-238B-1 (1-2130)

```

```

Qy 7 ValLeuValPheLeuLeuAlaGlyLeuProLeuGlnAlaLys-----Arg 24
|||
Db 46 CTCCTCATTTGGCTGTGATAGTGCTTTGGCTGTGGGGCTACAAAGTACCCAGA 105
|||||
Qy 25 PheArgValLeuGlyHisGluGlnIleTyProAspHisMetArgGluAsnGlnLeu 44
|||||
Db 106 AACCAGGACTGGCTTGGT-----GTCTCAAGGCMACTC 138
|||||
Qy 45 ArgGlyTrpSerSerAspGluAsnIleTrpAspGluLeuTyProValTrpArgArg 64
|||||
Db 139 AGA-----ACCAAGCCCTGGACAGCCAGCTGTCCAGAGTGGACA--- 180
|||||
Qy 65 GlnGlyValArgTrpLysAspSerTrpGlnGlyGlyArgValGlnAlaLeuThrSer 84
|||||
Db 181 ---GAAGCCACAGACTGTGCTGTGGAGAGGTGGTCAAGTGTCCCTCAAGTTCAGTA 237
|||||
Qy 85 AspSerProAlaLeuValGlySerAsnIlePheValValAsnLeuValPheProArg 104
|||||
Db 238 GNTGGGCTACAGTATGGGAGTGGCAATGCTCTCTCTATTTGGCTGTACTCTTGA 297
|||||
Qy 105 CysGlnLysGlnAspAlaAsnGlyAsnIleValTyGluArgAsn---CysArgSer 123
|||||
Db 298 AGCCAAAAGGTATTGGCAGATGGGAGGTATCTGGGTCAACATACCATCATCAATGG 357
|||||
Qy 124 LeuGluLeuAlaSerAspProTyValTyrrAsnTrpThrGlyAlaAspAspGluAsp 143
|||||
Db 136 AGCCAGGTGTGGGAGGACAGCCAGTGTATCCCAAGNACT-----GACAGTCC--- 408
|||||
Qy 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163
|||||
Db 409 -----TGCATCTTCCTGTATGTTGGGACTTTGGCCA 438
|||||

```

```

Qy 164 ArgProHisGlyArgLysLysTrpAsnPheValTyValPheHisThrLeuGlyGlnTr 183
|||||
Db 439 TCTGGCTCTGTGTCAGAGAGAGAGCTTTGTTATGTCGAGAGACTTGGGGCAATAC 298
|||||
Qy 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsnLeuThrVal 403
|||||
Db 499 TGGCAAGTTCTTGGGGGCCAGTGTCTGGCTGAGCATTTGGCAGCAGCAGGAGCAATGCTG 558
|||||
Qy 204 GlyProGlnValMetGlnValIleValPheArgHisGly---ArgAlaTyrrIlePro 222
|||||
Db 559 GGCACACACATAGAGTGTCTTCCATCGCGGGGATCCCGAGCTATGTGCT 618
|||||
Qy 223 IleSerLysValLysAspValTyValIleThrAspGlnIlePheValThrMet 242
|||||
Db 619 CTTCCTCATTCAGCTCAGCTCAGCTCCTACCATCTACTGACCATGCTTCTCGCTGAGCGTG 678
|||||
Qy 243 TyrGlnGlnAspAspArgSerSerAspGluThrPheLeuArgPheProIlePhe 268
|||||
Db 679 TCCGTTTGGGGGCTTGGATGAGGGGAGACAGCACTTCTGAGAAATCAGCTCTGCACC 732
|||||
Qy 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrrSerAlaIleSerTylys 282
|||||
Db 739 TTTGCCCTCCAGCTCCATGACCCAGTGTCTGTGCTGAAGCTGACCTCTCTCTACACC 798
|||||
Qy 283 TrpAspPheGlyAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
|||||
Db 799 TGGGACTTTTGGAGACAGTAGTGGAACTGTCTCTGGGCACTTGTGTGCTCATCTACT 858
|||||
Qy 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValIleThrAlaValPro----- 320
|||||
Db 859 TACCTGGAGGCTGGCGCCATCATCTGCCAGGTGTGCTGCAGGCTGCCATTCCTCTCACC 918
|||||
Qy 321 -----GlyProCysPro----- 324
|||||
Qy 919 TCTGTGGTCTCTCCCACTTCCAGGCCACACAGATGGGCACAGGCCAAGCTGACAGGCC 978
|||||
Qy 324 ----- 324
|||||
Db 979 CTTACACACAGCTGGCCAGTGGCTACTACAGAGTTTGGGTACTACAGCTGGTCAG 1038
|||||
Qy 325 SerProThrProSerProSerSerSerThr----- 334
|||||
Db 1039 GCGCAACTTCGACAGCCCTTGGAAACACATCTGTGAGTGGCAACCACTGAAGTCATA 1098
|||||
Qy 335 SerProSerProAlaSerSerProSerProThrLeuSerThrProSerProSerLeuMet 354
|||||
Db 1099 AGCACTGCACCTGTGCAGATSCCACTGACAGAGACAGCATGATGACACCTGAGAAGGTG 1158
|||||
Qy 355 Pro-----ThrGly 357
|||||
Db 1159 CCAGTTTCAGAGCTGTGGTGTACCACTGGCAGAGATGCAACTCCAGAGGCTTCAGGT 1218
|||||
Qy 358 TyrLysSerMetGlnGluSerAspIle----- 366
|||||
Db 1219 ATGACACCTGCAGAGGTATCAATTGTGTCTTCTGGNACCACAGCTGCACAGGTAAACA 1278
|||||
Qy 366 ----- 366
|||||
Db 1279 ACTACAGAGTGGTGGAGACCAACAGCTAGAGAGCTTACCTATCCCTGAGCTGAGAGTCCA 1336
|||||
Qy 367 -----SerAsnGlu----- 369
|||||
Db 1339 GATGCCAGCTCAATCATGTCTACGGAAGATATTACAGTTCCCTGGGCCCCCTCGCTGGAT 1398
|||||
Qy 370 -----AsnCysArgTleAsnArg 375
|||||
Db 1399 GGTACAGCCACTTAAGCTGTGGTGAAGAGACAGTCCCCCTGGATGTGTCTGTATCCA 1458
|||||
Qy 376 TyrGlyTyrrPheArgAlaThrIleThrIleValAspGlyIleLeuGluValAsnIleTle 395
|||||
Db 1459 TATGTTCTCTTTTCCGCTCACCTGGACATTTGTCAGAGGTATTGGAAGTGCAGAGTCTGT 1518
|||||

```

```

Oy 396 GluValAlaaspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
Db 1519 CAGGCT-----GTGCGTCCGGTGAAGGGAT-----GCAATTGGCTGACT 1560

Oy 416 ValThrCysValgylAlaThrProThrGlnAlaCysThrIleIleSerAspProThrCys 435
Db 1561 GTGTCCTGCCAAGGCGCGCGTCCCAAGAGCATGCGATGATCTCATCGCCAGAGGTGC 1620

Oy 436 GlnIleAlaGlnAsnArgValCysSerProValAlaAlaAspGluLeuCysLeuLeuSer 455
Db 1621 CAGCCCGCTGCCAGCGCGTGTGCGAGCGTGTCTGCCACCGCCAGCGCTGCCAGTGTGT 1680

Oy 456 ValArgGlnAlaPhe---AsnGlySerGlyThrTyrCysValAsnPheIleThrLeuGlyAsp 474
Db 1681 CTGCACCGACTAGTCAAGGCTGCTGGGAGACATACGCTCAATGCTGTCTGGCTGAT 1740

Oy 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleThrGlyAsnLeuGly--- 493
Db 1741 ACCAACAGGCTGGCAGCTGGTCAGACCCAGCAGCTTATCATGCTGTGTCAGAACGCGCTT 1800

Oy 494 SerProLeuAsnArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
Db 1801 GGSCAGGTCGCGTATGTCGGCATCTG-----CTGGTGTGATGCT 1845

Oy 514 ThrMetValThrIleLeuLeuTyrIleCysLys 523
Db 1846 GTGTCCTTGCATCTCTCATATATAGCGC 1875

RESULT 13
US-09-862-260A-1
; Sequence 1, Application US/09862260A
; Patent No. US2002008221A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: HEPATITIC ANTI-MELANOMA COMPOUNDS
; FILE OF INVOICE: 1268B121020
; CURRENT APPLICATION NUMBER: US/09/862,260A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/208,955
; PRIOR FILING DATE: 2000-05-31
; PUBLICATION NUMBER: 2002/02467,877
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2130
; ORGANISM: Homo sapiens
US-09-862-260A-1

Alignment Scores:
Pred. No.: 156e-56 Length: 2130
Score: 627.50 Matches: 176
Identity: 29.8% Mismatches: 205
Basecall Similarity: 27.6% Indels: 173
Query Match: 20.50% Gaps: 19
DB:

US-09-943-075A-2 (1-572) x US-09-862-260A-1 (1-2130)
Oy 7 ValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGlnAlaLys-----Arg 24
Db 46 CTTCTTCATTGGCTGCTGATAGTGTCTTGTGGCTGTGGGCGCTACAAAGTACCAGA 105

Oy 25 PheArgAspValPheGlyHisGluInTyrProAspHisMetArgGluAsnAsnGlnLeu 44
Db 106 ACCGAGGACTGGCTTGGT-----GTCTCAGGCGCACTC 138

Oy 45 ArgGlyTrpSerSerAspGluAsnGluTyrAspGluLeuLeuTyrProValTrpArgArg 60
Db 139 AGA-----ACCAAGCGCTGGAAACAGCAGCGCTGTATCCAGAGTGGACA--- 180

Oy 65 GlyGluGlyArgTrpIleAspSerTrpGlnArgValGlnAlaLeuThrSer 84

```

```

Db 181 ---GAGCCCGAGAGCTGACTCGGAGAGGTGTCAGGTCCCTCAAGTCAAGTANT 237
Oy 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeuValPheProArg 104
Db 238 GATGGCGCTACACTGTGTGGTGCAAATGCCCTCTCTCTGCTTGAATCTCCCTGGA 297
Oy 105 CysGlnGlyGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
Db 298 AGCAGAAAGGATGGCCAGATGGCGAGGTATCTGGGTCAACATACCATCAATGAG 357
Oy 124 LeuGlnAlaLeuSerAspProTyrValTyrAsnTrpThrThrGlyAlaAspAspGluAsp 143
Db 358 AGCAGCGCTGGGGAGAGACGCCCTGTATGCCCGAGGAACT-----GAGCATGCC--- 408
Oy 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhe 163
Db 409 -----TGCACTCTCCCTGATGTGGTGGACCTGTGCCCA 438
Oy 164 ArgProHisGlyArgLysTyrAsnPheValTyrValPheIleHisThrLeuGlyGlnTyr 183
Db 439 TCTGGCTCTGGTCTCAGAGAGAGCTGTTTATGTCTGGAGACCTGGGGCCCATAC 498
Oy 184 PheGlnIleLeuGlyGlyGlySerAlaArgValSerIleAsnThrValAsnLeuThrVal 203
Db 499 TGGCAAGTCTAGGGGCCCGCATGTCTGGGCTGAGCATGGACAGCAGGCAATGCTG 558
Oy 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222
Db 559 GGCACACACACCATGAGACTGACTGTCTACCATCCCGGGGATCCCGGAGCTATGTGCT 618
Oy 223 IleSerTyrValLysAspValTyrValIleThrAspGlnIleProIlePheValThrMet 242
Db 619 CTTCTTCATTCCAGCTCAGCCTTACCCATTACTGCCAGGTGCTTCTCCGTGAGCTG 678
Oy 243 TyrIleLysAsnAspArgAsnSerSerAspGluThrPheLeuAsnAspLeuIlePhe 262
Db 679 TCCCACTTGGGGCGCTTGGATGGAGGACAAACGACTTCTCTGAGAAATCAGCCTGTGACC 738
Oy 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLys 282
Db 739 TTTTCCCTCCAGCTCCATGACCCAGTGGCTATCTGGCTGAGCTGACCTCTCTACACC 798
Oy 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
Db 799 TGGAGCTTGGAGACAGTAGTGAACCTGATCTCTCGGCGACTTGTGTGCTACTCATACT 858
Oy 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnAlaValPro----- 320
Db 859 TACTTGGAGCTGGCCCGCTCACTGCCAGTGTCTCTGACAGCTGCCATCTCTCTCACC 918
Oy 321 -----GlyProCysPro----- 324
Db 919 TCGTGTGGCTCTCCCGCTTCCAGGCCACACACATGGGACGCCACGCCACTGACGAGGCC 978
Oy 324 ----- 324
Db 979 CTTTACACACAGCTGGCCCAAGTGGCTACTACAGAGTTGTGGTACTACACTGCTGCAG 1038
Oy 325 SerProThrProSerProSerSerSerThr----- 334
Db 1039 GGGCAACTGCCAGCCTCTGCAACCCACATCTGCGAGTGGCCACCACTGAAGTCAATA 1098
Oy 335 SerProSerProAlaSerSerProSerProThrLeuSerThrProSerProSerLeuMet 354
Db 1099 AGCACTGCACTGTGTCAGATGCCCACTGCCAGAGACACAGATATGACACCTGGAAGGTG 1158
Oy 355 Pro-----ThrGly 357
Db 1159 CCAGTTTCAGAGTCATGGGTACCACACTGGCCAGAGATGTCACACTCCAGAGCTACAGGT 1218
Oy 358 TyrLysSerMetGluLeuSerAspIle----- 366

```

Db 1219 ATGACACATCCAGAGGTATCAATTCGTGGTCTTCTGACACACAGCTGCACAGGTAAACA 1278
 Qy 366 ----- 366
 Db 1279 ACTACAGAGTGGGAGGACACACAGCTACAGAGTACCTATCCCTGAGCTGAGAGGTCCA 1338
 Qy 367 ----- 369
 Db 1339 GATGCCAGCTCAATCATGCTCTACGAAAGTATTACAGAGTTCTGGGCGCCCTGCTGGAT 1398
 Qy 370 ----- 375
 Db 1399 GGTACGCCACCTTAGCTGGTGGAGACACAGTCCCTCGATATGTTCTGTATGATCCA 1458
 Qy 376 TCGlyTyrPheAspGAlaThrIleValAspGlyIleLeuAlaValAsnIleIle 395
 Db 1459 TATGGTTCCTTTTCCCTGACCTGTCAGCATATGCCAGGATATGAAAGTGGCGAGATCTG 1518
 Qy 396 GlnValAlaAspValProIleProThrLeuGlnProAspAsnSerLeuMetAspPheIle 415
 Db 1519 GAGCTT-----GTGGCGTCCGTGGGGGAT-----GCATTGTAGCTGACT 1560
 Qy 416 ValThrCysIleGlyAlaThrProThrGluAlaCysThrIleIleSerAspProThrCys 435
 Db 1561 GTGTCTGTCACAGCGGGCTGCCCAAGAAAGCCTGCATGGAGATCTCATCGCCAGGGTGC 1620
 Qy 436 GluIleAlaIshnAspValCysSerProValAlaValAspGluLeuLeuSer 1650
 Db 1621 CAGCCCTCGCCAGCGCTGTGCCAGCTGTGCTACCCAGCCGCCCTGCCAGCTGT 1680
 Qy 456 ValArgArgAlaPhe---AsnGlySerGlyTyrCysValAsnPheThrLeuGlyAsp 474
 Db 1681 CTGCACAGCTACGTAGGGTGGCTTCGGGGACATACCTCCATGTGCTCTGCTGAT 1740
 Qy 475 AspAlaSerLeuAlaLeuThrSerAlaLeuIleSerIleProGlyLysAspLeuGly--- 493
 Db 1741 ACCAAGCTGGCAGCTGGTCCAGCACCCAGCTTATCATGCTGGTCCAGAACGAGCGCT 1800
 Qy 494 SerProLeuArgThrValAsnGlyValLeuIleSerIleGlyCysLeuAlaMetPheVal 513
 Db 1801 GGGCAGGTTCCCTGATCGTGGGCATCTG-----CTGGTGTGATGGCT 1845
 Qy 514 ThrMetValThrIleLeuLeuTyrLysLys 523
 Db 1846 GTGGTCTCTGCATCTGATATATAGGCC 1875

RESULT 14
 US-10-047-539-3
 ; Sequence 3, Application US/10047539
 ; Patent No. US20020177547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MOLLING, KARIN
 ; APPLICANT: PAVLOVIC, JOVANA
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS FOR TREATING OR PREVENTING
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: VOS-27
 ; CURRENT APPLICATION NUMBER: US/10/047-539
 ; PRIOR APPLICATION DATE: 2002-01-15
 ; NUMBER OF SEQ. ID NOS: 16
 ; NUMBER OF SEQ. ID NOS: 13
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ. ID NO 3
 ; LENGTH: 2131
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PROJECT: PHARMACEUTICAL
 ; NAME/KEY: CDS
 ; LOCATION: (12)..(2018)
 US-10-047-539-3

Pred. No.: 5,99e-56 Length: 2131
 Score: 622.00 Matches: 177
 Percent Similarity: 41.55% Conservative: 96
 Best Local Similarity: 26.94% Mismatches: 204
 Query Match: 20.32% Indels: 180
 DB: Gaps: 20

US-09-943-075A-2 (1-572) x US-10-047-539-3 (1-2131)

Qy 7 ValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGlnAlaAlaLys-----Arg 24
 Db 36 CTTCTTATTTGGCTGTGATAGGTGCTTCTGCTGTGGGGCTACAAAGATACCCAGA 95
 Qy 25 PheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgLysAsnGlnLeu 44
 Db 96 AACCAGGACTGGCTTGGT-----GTCTCAAGGCAATCTC 128
 Qy 45 ArgGlyTyrSerAspAspGluAsnGlnTyrAspGluGlnLeuTyrProValTyrArgArg 64
 Db 129 AGA-----ACCAAGCTTGAGACAGGACGTGTCCAGATGACAA----- 170
 Qy 65 GlyGluGlyArgTyrPlysAspSerTTPGluGlyArgValGlnAlaAlaLeuThrSer 84
 Db 171 ---GAAGCCAGAGACTATGCTCTCGAGAGGTGGTCAAGTGTCCCTCAAGTCAAGTAAT 227
 Qy 85 AspSerProAlaLeuValGlySerAsnIleThrPheValValAspAspLeuValPheProArg 104
 Db 228 GATGGCCTACTGATGTGTGCANAATGCTCTCTCTATTGCTGTACTTCCCTGGA 287
 Qy 105 CysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsn---CysArgSerAsp 123
 Db 288 AGCCAAAGGTATPCCAGATGGCGAGGTATCTGGGTCAACATACATCATCATATGGG 347
 Qy 124 LeuGluLeuAlaSerAspProTyrValTyrAsnTyrThrGlyValAspAspGluAsp 143
 Db 348 AGCCAGGTGTGGGAGGACAGCAGTGTATCCCGAGGAAT-----GACATGACC--- 398
 Qy 144 TrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLysProPhePro 163
 Db 399 -----TGCATCTCTCTCTGTGTGGACTTGGCCCA 428
 Qy 164 ArgProHisGlyArgLysLysTyrPAsnPheValTyrValPheHisThrLeuGlyGlnTyr 183
 Db 429 TCTGGCTCTTGGTCTCAGAGAAGAGCTTTGTTATGTCTGGAAGACTGGGGCCCAATAC 488
 Qy 184 PheGlnLysLeuGlyGlnCysSerAlaArgValSerIleThrValAsnLeuThrVal 203
 Db 489 TGGCAAGTCTAGGGGCCAGCTGCTGGCTGAGCATTTGGGACAGAGAGGCAATGCTG 548
 Qy 204 GlyProGlnValMetGluValIleValPheArgArgHisGly---ArgAlaTyrIlePro 222
 Db 549 GGCACACACACATGAAGTGTACCATCGCGGGGTCCCGGAGCTGTGTGCTT 608
 Qy 223 ThrSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheThrMet 242
 Db 609 CTTGCTCATTCAGCTCAGCTTCCACCATCTACTGACCAGGTGGCTTCTCTCGTGAAGCTG 668
 Qy 243 TyrGlnLysAsnArgAsnSerSerAspGluThrPheLeuArgAspLeuProIlePhe 262
 Db 669 TCCAGTTCGGGGCTTGGTGGAGGACACACACTCTCTGGAATACAGCTCTTGACC 728
 Qy 263 PheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSerTyrLys 282
 Db 729 TTTGGCTCTCAGCTCCTGACCTAGTGTGCTATCTGCTGAAGCTGACCTCTCTACACC 788
 Qy 283 TrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsnHisThr 302
 Db 789 TGGGACTTTGGAGNCAGTAGTGGAGCCCTGATCTCTCGGCGACCTGTGGTCACTACT 848
 Qy 303 TyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro----- 320
 Db 849 TACCTGAGCTGGGCCAGCTCATCTGCCCGAGGTGGTCTGCGAGGTGCCATCTCTCTCACC 908

Alignment Scores:


```

QY 519 LeuLeuTyrLysLysHisLysLysTyrLysProIleGlyAsnCysThrArgAsnValVal 538
      ||||| : : : : : |||||
Db 1744 ATATAGGCATAGACTTAAGACGACGGCTCAGTTTCCCAATGCCACATG----- 1794
QY 539 LysGlyLysGlyLeuSerValPheLeuLeuSerIleSAlaLysAlaProPheSerArgGlyAsp 558
      : : : : : : : : : |||||
Db 1795 -----GTAGCACTACTGCTGGCGTGCCTCGCGTCTCGCGCGCGCGGCGCTT 1842
QY 559 ArgGlnLysAspProIleuLeu 565
      |||||
Db 1843 GGAGAAACAGCCCGCTCTC 1863

Search completed: April 9, 2003, 16:00:55
Job Time: 181 secs

```

Search completed: April 9, 2003, 16:00:55
Job time : 181 secs